مقدمة

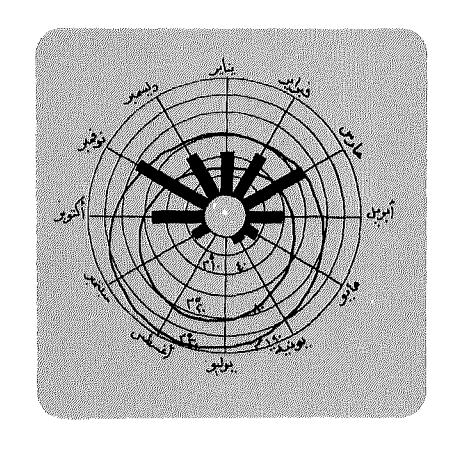
الأساليب الكمية فى الجغرافيا

الأستاذ الدكتور

فتحى عبد العزيز أبو راضى

أستاذ الجغرافيا الطبيعية

وعميد كلية الآداب، جامعة بيروت العربية

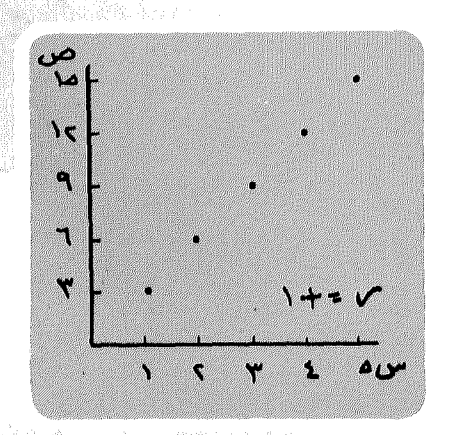


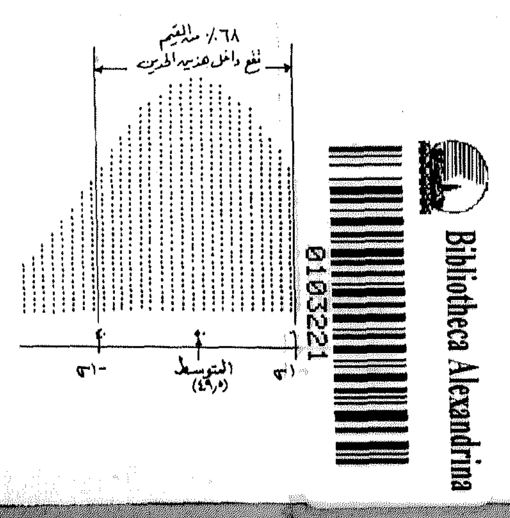
-أ) ك١ + (س٣ - أ)

ك١ + ك٢ + ك٣ ..

ك٣ ... + (س ن)

..... ك ن





دار المعرفة الجامعية





مقدمة

الأساليب الكمية فى الجغرافيا

دكتور

فتحى عبد العزيز أبو راضى

أستاذ الجغرافية الطبيعية

عميد كلية الآداب

جامعة بيروت العربية

٢٠٠٠

دار المعرفة الجامعية

٤٠ ش سوتير - الأزاريطة - ت ٤٨٣٠١٦٣

٣٨٧ ش قنال السويس - الشاطبى - ت ٥٩٧٣١٤٦

دار المعرفة الجامعية

للطبع والنشر والتوزيع

* الإدارة : ٤٠ شارع سوتير
الأزاريطة - الاسكندرية
ت : ٤٨٣٠١٦٣

* الفرع : ٣٨٧ شارع قنال السويس
الشاطبى - الاسكندرية
ت : ٥٩٧٣١٤٦

بسم الله الرحمن الرحيم



إهـــداء

إلى زوجتي:

شريكة الحياة ....

وصديقة العمر ....

التصدير

شهدت سنوات العقود الثلاثة الماضية تغيراً كبيراً وتطوراً ملحوظاً فى علم الجغرافيا، ليس فى منهجه ومحتواه فحسب، وإنما أيضاً فى الأساليب التى يعتمد عليها فى تحقيق أغراضه وأهدافه. وليس من شك فى أن أهم التطورات التى شهدها هذا العلم هو التعامل مع الأرقام أو ما عرف بالاتجاه الكمى المتمثل فى تطبيق الأساليب الاحصائية الكمية فى تحليل العلاقات المختلفة بين مكونات البيئة ونشاط الإنسان، وفى دراسة المشكلات والظاهرات الجغرافية المتنوعة سواء كانت طبيعية أو بشرية بغية الوصول إلى نتائج رقمية محدودة تختصر كثيراً من التحليلات الوصفية الكيفية (النوعية) فى تشخيص وتفسير الظاهرات الجغرافية.

ورغم أن لغة الأرقام ليست بغريبة على بعض فروع علم الجغرافيا، إذ تعد البيانات الرقمية مصدراً رئيسياً تعتمد عليه مجالات التحليل وأساساً للبحث والدراسة منذ البداية فى فروع مثل الجغرافية الاقتصادية وجغرافية السكان والجغرافية المناخية، فإن ما شهدته السنوات الأخيرة يعد محاولة للتعمق فى التعامل مع الأرقام وتحولاً نحو استخدام الأساليب الكمية فى التحليل الرقمى فى الدراسات والبحوث الجغرافية بعامة. وكان لهذا التطور أو التحول نتائج هامة أسفرت عن دفع عجلة هذا العلم والارتقاء به إلى مصاف العلوم الأرضية والاجتماعية التى اتخذت لنفسها

مساراً ديناميكياً كمياً. ولذا فقد أطلق البعض على هذا التطور أو التحول فى استخدام أسلوب التحليل الرقمى فى الجغرافيا إسم الثورة الكمية (Burton, 1963) والثورة المفاهيمية (Davis, 1972). وقد لقى ذلك ترحيباً من بعض الجغرافيين، وتحمساً، وأحياناً تعصباً، من البعض الآخر، ولكن بعد أن هدأت «الثورة» ظهر إسم «الجغرافية الكمية Quantitative Geography كمصطلح شامل لكل ما حدث من تطور وتحول فى الجغرافيا، ولذا فلا غرو أن يكون هو المصطلح السائد بين الجغرافيين المعاصرين. والجغرافية الكمية إذن ليست فرعاً من الفروع المعروفة للجغرافيا، كما أنها ليست فرعاً «جديداً» New Geography، وإنما هي تشير إلى واقع جديد أخذ يفرض نفسه على الجغرافيا، أو بعبارة أخرى هى اتجاه جديد فى معالجة المشكلات الجغرافية بعامة. ويتمثل هذا الواقع أو الاتجاه فى مجموعة الأساليب الاحصائية الكمية التى كان لكثرة تطبيقها واتساع مجال استخدامها وتعدد الدراسات فيها أن أدخلت قواعد وقوانين جديدة على الفكر الجغرافى، كما مهدت اكتشاف آفاق جديدة لميدان الدراسات الجغرافية. وكان من نتيجة ذلك أن تزايد الاهتمام الكمى وتطور فى مختلف الفروع الجغرافية، وليس أدل على ذلك من أنه لاتخلوا الآن المناهج الدراسية للجغرافيا فى أقسام الجغرافيا بجامعات العالم المتقدم من منهج فى «الجغرافية الكمية».

ويهدف هذا الكتاب إلى محاولة تطبيق بعض الأساليب الكمية فى الجغرافيا واستخدامها كأدوات لحل المشكلات واستخلاص النتائج لاتخاذ القرارات التى هى أساس وهدف البحث العلمى النهائى. ومعنى ذلك أن هذا الكتاب لن يتناول دراسة وتطبيق كل الأساليب الممكنة فى هذا المجال بل سيقتصر على «عينة» من «مجتمع» الأساليب الكمية الذى مازالت الدراسة فيه ميداناً متسعاً تستقى منه كثير من الموضوعات فى مجال الدراسات الكمية المتقدمة.

ولقد أعدت موضوعات هذا الكتاب، التى هى حصيلة جهد وخلاصة تجربة تدريس لهذا اللون من الدراسة على امتداد نحو ست عشرة سنة لطلاب الجغرافيا فى جامعات شفيلد ونوتنجهام - انجلترا - والاسكندرية. لكى تكون بمثابة مقدمة (مدخل) لدراسة الأساليب الكمية فى الجغرافيا، تستطيع أن تعطى الباحث المبتدئ

معرفة منظمة عن الأساليب والأدوات المستخدمة فى جمع وتحليل بيانات الظاهرة الجغرافية، كما يمكن أن تكون أساساً يعتمد عليه الباحث المتخصص فى تحليله للحقائق الرقية الخاصة بالمشاكل التى تتعرض لها دراسته كبغية الوصول إلى أقصى ما يمكن من نتائج عليه مفيدة. ولقد حرصنا فى ضوء هذا الهدف أن يكون عرض هذه الأساليب مبسطاً ولكنه وافياً وشاملاً للعديد من المبادئ التى تعد ضرورية وأساسية لمعالجة المشكلات الجغرافية. كما حاولنا بقدر المستطاع أن نتجنب الإثباتات والبراهين الرياضية حتى يتمكن الباحث الذى لايتمتع بأية خلفية رياضية من فهم واستيعاب هذه الأساليب ومتابعة تطبيقها دون عناء. وأياً كان الأمر فإن اختيار وعرض موضوعات مثل هذا الكتاب ليس بالأمر اليسير، إذ لاتوجد «طريقة» معروفة أو متفق عليها لتقسيم وتنظيم موضوعات هذا اللون من الدراسة، ولذا فإن التسلسل فى موضوعات (فصول) هذا الكتاب يتخذ من المشكلات المختلفة التى تواجه الباحث الجغرافى أساساً له مثل تلخيص البيانات والمعطيات الخاصة بموضع (حيز) واحد فى مكان (مجال أو فراغ) معين، أو توضيح المقارنات والعلاقات بين متغيرين أو أكثر. أو بمعنى آخر موضوع الكتاب بعامة يتجه نحو تحليل ومعالجة المشكلات الجغرافية أكثر من الاتجاه نحو دراسة الأساليب المستخدمة لذاتها وذلك على الرغم مما يوحى به أو يشير إليه، عنوان الكتاب «الأساليب الكمية». وحتى لايحدث بعض من اللبس فإننا قمنا بتخليص الخصائص الأساسية والأهداف الرئيسية المنشودة من الأسلوب المستخدم فى التحليل فى بعض المواضع، إذا لزم الأمر، بالإضافة إلى شرح أمثلة تطبيقية لمشكلات نوعية توافق to fit خصائص وأهداف الأسلوب المستخدم فى مواضع أخرى.

ويتكون إطار هذا الكتاب من مقدمة وأربعة أبواب تضم اثنتا عشر فصلاً وتتناول المقدمة شرح المفاهيم الاحصائية التى سيتردد ذكرها فى كثير من المواضع، وبيان الخصائص الرئيسية التى ينبغى توافرها فى البيانات الاحصائية التى تمثل المادة الخام لأسلوب التحليل الكمى. وعرضنا فى نهاية هذه المقدمة لموضوع الدقة والأخطاء فى البيانات وأسبابها. ويعرض الباب الأول والذى اشتمل على الفصول الثلاثة الأولى لطرق جمع البيانات وإعدادها للتحليل الكمى. وقد اشتمل هذا الباب على مناقشة تحليلية مختصرة للمبادئ الأساسية لأساليب تحديد حجم العينات المتعارف

عليها، مع ربط هذه المبادئ بالواقع الجغرافي. وقد قمنا في هذا الباب أيضاً بدراسة طرق العرض البياني والجدولي كمنطلق أساسي للتحليل الاحصائي الكمي.

ويختص الباب الثاني يتكون من ثلاثة فصول، من الفصل الرابع حتى الفصل السادس، بمناقشة موضوع الوصف الكمي الذى يعتبر عملاً أساسياً في كل العلوم - بما فيها الجغرافيا - عن طريق استخدام مقاييس محددة ومعايير ثابتة متفق عليها، وقد عرضنا في هذا الباب مقاييس النزعة المركزية (المتوسطات) وبينا كيفية تطبيقها والفروق في النتائج المستخلصة من كل منها. وبعد ذلك انتقلنا إلى مناقشة مقاييس تجانس واختلاف (تشتت) البيانات وانحرافها (تباعدها) عن قيمة أحد مقاييس المتوسطات وبينا كيفية استخدامها وخصائصها ومجال تطبيقاتها المتعددة إلى جانب عرض مفصل لمزايا استخدام المقاييس ومشاكل تطبيقاتها، وقد تناولنا في هذا الباب أيضاً موضوع اتجاه وشكل تركز بيانات المتغيرات في أحد نواحي توزيعها، أو ما يعرف بمؤشرات التركز، إذ لاتكفي مقاييس المتوسطات والاختلاف (التشتت) في وصف وتشخيص التوزيعات ومقارنتها بعضها البعض لتحديد خصائصها وملامحها. ويشتمل الباب الثالث على أربعة فصول، من السابع حتى العاشر، تختص باستعراض التقدير الاحصائي وأساليب المقارنة بين مجموعات البيانات. وقد ناقشنا وبينا طرق استنتاج وتقدير خصائص المجتمع من البيانات التي جمعت بأسلوب المعاينة بسبب عدم إمكانية دراسة وفحص جميع مفردات المجتمع التي جمعت بأسلوب المعاينة بسبب عدم إمكانية دراسة وفحص جميع مفردات المجتمع التي قد تكون كثيرة جداً. وانتقلنا بالمناقشة في هذا الباب إلى قواعد إختيار الفروض الاحصائية التي تعتمد عليها أساليب المقارنة بين البيانات. وبينا أساليب معالجة البيانات ومقارنتها في صورة عملية تطبيقية على البيانات الجغرافية للوقوف على أهمية الأساليب المعلمية (البارامترية) والأساليب غير المعلمية (غير البارامترية) التي تستخدم هذا الشأن.

ويخص الباب الأخير من الكتاب وهو الباب الرابع باستعراض وتحليل العلاقات بين المتغيرات وتحديد اتجاهها العام. ويشتمل هذا الباب على ثلاثة فصول من الحادى عشر حتى الثالث عشر. وقد خصص الفصل الأول منها لتحليل الارتباط

لمعرفة درجة واتجاه العلاقة بين المتغيرات. والفصل الثانى من هذا الباب يعالج تحليل الانحدار لتمثيل العلاقة بين بيانات متغيرين بطريقة رياضية يمكن التنبؤ بها والحصول منها على بيانات متغير معين كلما تغير الآخر. وبهذا يمكن تقدير سلوك أحد المتغيرات فى ضوء تأثره بمتغير آخر أو بعدة متغيرات أخرى.

والفصل الثالث من هذا الباب يتناول بالشرح والتفصيل تحديد الاتجاه العام الذى يعكس تأثير العوامل المختلفة التى تؤدى إلى التغيرات أو التطورات فى الظواهر كمياً عبر الزمن. وقد بينا أن استمرار الاتجاه العام فى المستقبل هو الأساس المنطقى للتنبؤ الاحصائى الذى يعتمد عليه عند اتخاذ القرارات فيما يتعلق بتخطيط المستقبل، والهدف الرئيسى لتحليل «السلاسل الزمنية» عن طريق دراسة مركباتها الأساسية وأهميتها فى استنتاج التذبذبات وتتبعها فى منحى الظاهرة.

وحتى تكتمل الفائدة العلمية من موضوعات الكتاب، فقد حرصنا على تزويده بأمثلة عديدة وتطبيقات كثيرة لبيانات جغرافية لتوضيح استخدام وتطبيق أدوات وأساليب التحليل الاحصائى الكمى وتفسير نتائجها، بالإضافة إلى الكثير من الأشكال والرسوم البيانية حتى تكتمل متابعة الحقائق الواردة فى متن الموضوعات، كما لم يفتنا إلحاق مجموعة من الجداول الاحصائية التى يستعان بها فى عملية المعايرة الاحصائية.

ولاندعى أن الكتاب يخلو من نقائص فليس فى وسع أى باحث مهما كانت مقدرته العلمية أن يصل بدراسة إلى درجة الكمال - فهو لله وحده - ولكنها محاولة أرجو من خلالها أن أكون قد حققت ولو إضافة بسيطة إلى المكتبة العربية التى تعانى النقص الشديد من مؤلفات فى الأساليب الكمية فى مجال البحوث والدراسات الجغرافية بخاصة. كما أن كل أملى أن تحقق هذه المحاولة الهدف المنشود منها، وتسد هى وأمثالها الفجوة العلمية التى تفصل بين البحث الجغرافى فى الجامعات العربية وبقية جامعات العالم المتقدم، وأن يكون «تباين» المعروض من الموضوعات واتساع وتعمق الدراسة فيه خير معين للباحثين والمهتمين بالدراسات الجغرافية المتطورة، ومشجع لهم على تفهم طبيعة وخصائص الأساليب الكمية والتعامل معها فى بحوثهم ودراساتهم ففى ذلك مواصلة السير فى نهج المعرفة المتطورة ومواكبة التقدم العلمى الخلاق.

وأود بهذه المناسبة أن أتقدم بالشكر الجزيل لأستاذى الجليل الأستاذ الدكتور جودة حسنين جودة الذى كان لتوجيهاته السديدة القيمة ولتشجيعه المتواصل لى أثر كبير دفعنى نحو إصدار هذا الكتاب. كما أتوجه بجزيل الشكر وعظيم الامتنان إلى أساتذتى الأفاضل وزملائى فى أقسام الجغرافيا بجامعة الإسكندرية وجامعتى شفيلد ونوتنجهام بانجلترا على ما قدموه لى من عن صادق وتشجيع دائم خلال مرحلة دراستى وتدريسى لهذا الموضوع، والذين أفدت كثيراً من توجيهاتهم وملاحظاتهم أثناء مرحلة إعداد هذا الكتاب.

ويبقى أن أرجو بعد العناء أن أكون قد وفقت إلى أن أوفى فيما أقصد إليه على غاية، وأن يحقق هذا الكتاب الغرض من اصداره.

والله الموفق والمستعان ،،

دكتور فتحى عبد العزيز أبوراضى

الإسكندرية فى ١٠ / ٨/ ١٩٩٩

مقدمة
فى
المفاهيم الإحصائية

مقدمة
فى
المفاهيم الإحصائية

تحتل – الآن – أساليب التحليل الإحصائى الكمى أو «الطرق الإحصائية» كأساليب علمية وأدوات بحث، أهمية خاصة فى الأبحاث العلمية الحديثة؛ إذ لايكاد يخلو أى بحث من دراسة تحليلية إحصائية تتعرض لأصل الظاهرة قيد البحث فتصور واقعها فى قالب قياسى رقمى وتنتهى إلى إبراز اتجاهاتها وعلاقتها بالظاهرات الأخرى، كما لاتوجد أى دراسة دون أن تكون قد اتخذت لنفسها – أو لأحد جوانبها على الأقل – مناراً كمياً قائماً على الأساليب الإحصائية والرياضية الحديثة ونظرياتها المتجددة ونماذجها المتطورة، وعلى أحدث الآلات الحاسبة وأسرعها. ولايقتصر هذا الاتجاه، الذى يشير إلى واقع جديد وتطور حديث، فى طريقة دراسة المشكلات ومعالجة البيانات على فرع من فروع العلم بل تتبناه جميع الفروع تقريباً للاستفادة منه، حتى أصبح استخدام الإحصاء أو الطرق الإحصائية سمة رئيسية وضرورة ملحة للبحث العلمى.

ويشير مصطلح «إحصاء» إلى معنيين رئيسيين (Hays, 1970): فهو فى معناه الضيق يستخدم للتعبير عن الكميات الهائلة من البيانات Data التى تجمع عن

طريق الإستفتاء أو التجارب أو الحصر مثل الإحصاء السكانى، كميات الإنتاج الزراعى والصناعى، وحجم التبادل التجارى بين الدول. أى أنه بهذا الإستخدام يختص بالحقائق والأرقام Facts and Figures. أما المعنى الثانى فيختص بالطرق العملية والأساليب العلمية (البنيطة والمعقدة) فى معالجة وتحليل وتفسير البيانات بغرض الوصول إلى نتائج مقبولة وقرارات سليمة. والإحصاء بمعناه الأخير يمكن اعتباره أحد فروع الرياضة التطبيقية بما يحمل بين طياته من نظريات وقوانين تساهم بدرجة كبيرة فى اتخاذ القرارات التى أصبحت الأساس والهدف النهائى للبحث العلمى الخلاق.

والإحصاء كعلم يعتبر من العلوم الحديثة نسبياً. إذ بدأ يطلّ على العلوم الأخرى ويتصل بها فى أواخر القرن التاسع عشر. ولو أن لهذا الاتصال جذور أو ربما بذور وضعت خلال القرنين السابع عشر والثامن عشر حين اهتم علماء الرياضيات بوضع نظرية الاحتمالات (Hammond and Mccullagh, 1974). ففى سنة ١٧٧٠ استخدم مصطلح «الإحصاء Statistics» لأول مرة، ولكن بمعنى يختلف عن معناه فى الوقت الحاضر. فقد احتوى كتاب هوبر Hooper الذى نشر فى عام ١٧٧٠ بعنوان: "The Elements of Universal Erudition" فصلاً بعنوان «الإحصاء، Statistics» يتضمن تعريفاً للإحصاء على أنه أحد فروع علم السياسة الذى يهتم بجمع وتصنيف وتحليل الحقائق التى لها تأثير على حالة الدول المعروفة فى ذلك الوقت (Yale and Kendall, 1953). وفى أوائل القرن التاسع عشر حدث تغير (تحول) فى مفهوم علم الإحصاء، فأصبح يعرف على أنه علم «الحساب السياسى» الذى يستخدم الطرق العددية Numerical Methods فى إبراز الشخصية الاقتصادية للدولة. وبنهاية القرن الماضى استخدمت الطرق الاحصائية كأداة لتلخيص البيانات ووصفها فقط. ثم زاد الاهتمام بعلم الإحصاء مع بداية القرن الحالى حتى أصبح الآن علماً له قواعده ونظرياته، وأصبحت الأساليب الاحصائية تطبق كوسيلة لاستنتاج الحقائق فى شتى فروع العلوم الأرضية Earth Sciences

سواء كان فى العلوم الطبيعية كالكيمياء والطبيعة أو فى العلوم الاجتماعية كالجغرافيا والاجتماع وعلم النفس وغيرها.

ونظراً لأن مجال الدراسة التى بين أيدينا لاتمكننا من التعرف - بالتفصيل والتدقيق - على التطور التاريخى لعلم الإحصاء، تبعاً لتعدد طرق ومداخل دراسة هذا التطور، إلا أنه قد يكون من المفيد والمناسب هنا الإلمام بخواص هذا العلم عن طريق عرض مختصر لأهم وظائفه واستخداماته والمفاهيم الخاصة به.

الوصف الاحصائى Statistical Description

تعد وظيفة الوصف من الوظائف الأولية لعلم الإحصاء التى تستخدم فى تلمس حقائق الظاهرات المختلفة (جغرافية، اجتماعية، اقتصادية ... إلخ). ويعتمد أسلوب الوصف الإحصائى على المقاييس والمؤشرات الإحصائية، بعضها خاص بقياس القيم المركزية للبيانات. والبعض الآخر خاص بقياس مدى دقة هذه المقاييس. بغرض تقصى الحقائق وتحديد الخصائص العامة دون الوصول إلى نتائج أو استدلالات. وباستخدام هذا الأسلوب فى وصف الظاهرات بعد تلخيصها وعرضها على هيئة جداول أو أشكال بيانية يصبح من السهل تحديد خصائصها واتجاهاتها العامة بطريقة علمية منظمة.

الاستدلال (الاستنتاج) الاحصائى Statistical Inference

يعرف الإستدلال الإحصائى بأنه عبارة عن التوصيات أو الاستنتاجات المبنية على طرق إحصائية تتناسب مع ظروف الظاهرة موضع الدراسة. وتتناول تحليل بياناتها. ويمكن تقسيم الاستدلال الإحصائى إلى نوعين رئيسيين (Spiegel, 1972) النوع الأول يطلق عليه الإستدلال الإستنتاجى Deductive Inerence، وهو عبارة عن تطبيق تفسير ظاهرة خاصة من نتائج دراسة ظاهرة عامة. فعلى سبيل المثال إذا ذكرنا أن انخفاض نسبة وفيات الأطفال عموماً يرجع إلى ارتفاع المستوى الصحى

(ظاهرة عامة)، فإنه يمكن القول أن ارتفاع المستوى الصحى لسكان منطقة ما سيؤدى، بالضرورة، إلى انخفاض نسبة وفيات الأطفال بها. أما النوع الثانى من الاستدلال فهو الاستدلال الإستقرائى Inductive Inference، وهو تطبيق تفسير ظاهرة خاصة على ظاهرة عامة. وهذا النوع من الإستدلال يستخدم فى إطار التجارب المعملية، حيث أنه يمكن التحكم فى جميع الظروف المحيطة والمؤثرة فى الظاهرة موضع الدراسة. ولكن تحت الظروف الطبيعية فإن أى ظاهرة عرضة للاختلاف والتباين تبعاً لتأثير العوامل البيئية المتعددة، وعليه فلا يكون الاستدلال مؤكداً، وحيث ما هو صحيح بالنسبة للظاهرة الخاصة قد لايكون أن صحيحاً بالنسبة للظاهرة العامة، فإن لغة الإحتمال تستخدم عند عرض النتائج والتوصيات. فمثلاً تفسير ظاهرة التعرية الساحلية فى منطقة ما لاينطبق بالضرورة على نفس الظاهرة فى كل جهات العالم الساحلية. إذ أن العوامل الطبيعية التى تسبب هذه الظاهرة تختلف من منطقة إلى أخرى حسب الظروف البيئية السائدة فى كل منطقة. ولكن يمكن القول بأنه من المحتمل أن تشترك جهات العالم الساحلية التى تعانى من ظاهرة التعرية فى بعض العوامل المسببة. أى أن الاستقراء يعبر عنه فى هذه الحالة فى صورة احتمالية.

اختبار الفروض الاحصائية Testing of Hypotheses

عند تحليل البيانات الاحصائية يقوم الباحث بتلخيص الغرض من الدراسة وذلك على هيئة فرض مقترح أو نظرية، تتناسب مع خواص متغيرات ظاهرة، يمكن اختبارها والحكم على صلاحيتها لتفسير مدى إنطباق الفرض الموضوع على النتائج المتحصل عليها. فإذا كانت النتائج تتفق مع الفرض المقترح فإننا نقبله، وبالتالى يمكن تعميمه. أما إذا ظهر تعارض بين النتائج المستخلصة والفرض المقترح، فإن هذا الفرض سوف يرفض. ويتم قبول أو رفض الفرض المقترح

باستخدام الأساليب الاحصائية التى يتم بواسطتها تحليل البيانات الاحصائية للظاهرة موضع الدراسة، والتى عن طريقها يمكن الوصول إلى النتيجة العملية والقرار المناسب فى هذا الشأن.

ويتم الأسلوب الإحصائى لاختبار الفروض Testing of Hypotheses من خلال المشاهدة المتكررة للتغير فى الظاهرة، وعلاقة هذا التغير بالفرض المقترح، أو ما يسمى احصائياً بفرض العدم Null Hypothesis. فإذا ما توصلنا إلى عدم وجود فرق جوهرى أو حقيقى بين المشاهدات وما تم افتراضه، فإن الفرض المقترح يكون صحيحاً إحصائياً فى حدود خطأ مسموح به عند مستوى دلالة معين. وفى حالة توصلنا إلى وجود فرق حقيقى (معنوى) بين ما تم قياسه من واقع المشاهدات وما تم افتراضه فإن الفرق يكون غير صحيح، لأن المشاهدات الواقعية لاتؤيد ما كان يتوقع فى تغير بيانات الظاهرة موضع التحليل.

وتعتبر الاختبارات الاحصائية للفروض بمثابة الأسلوب العلمى فى استخلاص النتائج بطريقة موضوعية دقيقة. كما أنها مفيدة جداً فى شرح وتوضيح بعض الحقائق عن طبيعة البيانات الاحصائية.

التنبؤ (التوقع) الاحصائى: Statistical Prediction

يقصد بالتنبؤ، كمفهوم احصائى، هو تلك التغيرات التى حدثت لظاهرة ما فى الماضى، وليس فى المستقبل. وذلك لتأكيد وجود الظاهرة من خلال المشاهدة والقياس، واختبار الفروض وتفسير التغيرات واستخلاص النتائج. وتعتمد دقة التنبؤ اعتماداً يكاد كلياً على مبدأ «الحتمية Determinism» فى الظاهرة موضع التنبؤ. والذى يؤدى إلى استخلاص نتائج متشابهة تحت ظروف متشابهة. ولنضرب مثالاً على ذلك بالجاذبية الأرضية. فمن المعروف أن سرعة أى جسم فى الفراغ ترجع إلى الجاذبية نحو الأرض التى قدرت بنحو ٩,٧٥ متر/ ثانية/ ثانية. وبواسطة هذا القانون يمكن لنا أن نحسب – أو نتنبأ بتأكيد تام – طول المسافة التى سيقطعها جسم ساقط فى وقت معلوم؛ أو بعبارة أخرى يمكن التنبؤ بسرعة هذا الجسم فى

لحظة معلومة خلال فترة سقوطه. وفي الجغرافيا لانجد سوى عدداً قليلاً من الظاهرات الجغرافية التي تتصف بالطبيعة الحتمية، أما الغالبية العظمى منها فتتصف بتأثر بعضها البعض بطرق متباينة، وفي أوقات مختلفة. فنادراً ما يمكن الحصول على نتائج نهائية في دراسة أى منها حتى ولو كان ذلك تحت ظروف معينة، أو بوضع شروط أو فروض محددة.

والتنبؤ Prediction بمفهومه الاستدلالي السابق هو تنبؤ يخص الماضي Post-diction، وهو يختلف عن تنبؤ المستقبل Forcasting الذى يستخدم فيه التحليل الإحصائي للتوصل إلى توضيح الاتجاه العام لما سيحدث في المستقبل بالمتغيرات التي تتحكم في تطور ظاهرة ما. وكذلك بيان العلاقات بين متغيرات الظاهرة لفترة مستقبلة.

المجتمع والعينة Population and Sample

يعرف المجتمع الاحصائي Population أحياناً باسم «العشيرة» وهو يتكون من جميع المفردات Individuals (القيم Values) موضع الاستقصاء، والمطلوب معرفة خصائصها وتحديد الحقائق عنها. مثال ذلك المجتمع السكاني في دولة ما، والمجتمع الحيواني والنباتي، وكذلك مجتمع المصانع أو مجتمع الوحدات المنتجة من مصانع سلعة معينة.

والمجتمع الاحصائي إما أن يكون محدوداً Finite، أى أنه يمكن حصر جميع أفراده مثال ذلك مجتمع السكان في مدينة ما في سن معين (١ – ٥ سنوات). أو إنتاج مصنع للسماد في يوم معين. وإما أن يكون غير محدود Infinite، ومن أمثلته مجتمع النقط التي يتكون منها خط، أو عدد القياسات لمساحة ما، أو المجتمع المكون من جميع النتائج الممكنة (صورة ، كتابة) في قذفات متتالية لعملة معدنية. وعلى العموم فإن المجتمع الاحصائي يعد غير محدود في حالة صعوبة حصر أو قياس المفردات التي يتكون منها. وعند دراسة المجتمعات الاحصائية لايجب تطبيق القوانين الخاصة بالمجتمعات غير المحدودة على المجتمعات المحدودة. إلا إذا كانت الأخيرة مكونة من مفردات يزيد عددها عن عدة مئات أو آلاف.

وتعرف المقاييس الإحصائية الخاصة بالمجتمع والمميز له باسم معالم (ثوابت)

المجتمع Parameters مثل المتوسطات والانحراف المعيارى للمجتمع. وكما ذكرنا، فإنه قد يكون من المستحيل حساب هذه المعالم من المجتمع لصعوبة حصر جميع مفرداته، إلا أنه يمكن أن نستخلصها عادة من عينة مأخوذة من نفس المجتمع.

والعينة Sample هى جزء صغير من مفردات المجتمع الاحصائى تؤخذ لتمثل المجتمع، وتدرس بهدف الحصول على نتائج مهمة عن المجتمع عن طريق تحليل بياناتها. لذلك يجب أن تكون العينة ممثلة تمثيلاً صحيحاً للمجتمع الذى أخذت منه.

وتعرف المقاييس المحسوبة من بيانات العينة باسم «احصائيات» أو مقاييس العينة Statistics وهى مقاييس تقديرية Estimate لما يقابلها من «معالم» المجتمع التى تمثلها.

البيانات (المعطيات) الاحصائية Data

يقصد بتعبير البيانات أى «كمية من المعلومات فى صورة رقمية». والصورة الرقمية للبيانات تبدو إما على شكل أرقام صحيحة Integers مثل ١٠، ١١٢، ٤٦٤ ... إلخ، أو على شكل أرقام حقيقية Real Numters مثل ٢٠,٤، ٦١,٨، ١٨٢,١، ٣١٧,٠ أى أنها الأرقام التى تحتوى على علامة عشرية.

وتعد المعلومات الرقمية (البيانات الكمية) المادة الخام لأسلوب العمل الإحصائى، كما أنها تلعب دوراً كبيراً فى تطبيق أساليب التحليل الإحصائى الكمى حيث أن هذا النوع من البيانات يمكن قياسه كمياً مما يسهل استخدام هذه الأساليب لاستخلاص النتائج واتخاذ القرارات. وهناك بعض المصطلحات الخاصة بالبيانات الاحصائية التى سوف ترد كثيراً فى متن الفصول القادمة، نعرضها بصورة مختصرة قبل الخوض فى دراسة أساليب التحليل الإحصائى.

المفردات والمتغيرات Individuals and Variables

المفردة فى الإحصاء عبارة عن وحدة قياس المجتمع الإحصائى، ويعبر عن المفردات فى البيانات الاحصائية بالتمييز العددى للأشخاص كأعداد الطلبة والأسر والعمال، أو للحيوان مثل عدد الأبقار وعدد الأغنام، أو للجماد مثل عدد المصانع وعدد المدارس وعدد المستشفيات.

والمتغيرات Variables عبارة عن ظاهرات أو صفات تختلف قيمهما باختلاف الحالات Cases. والقيمة التى تعطى لكل مفردة من صفة معينة تعرف باسم Vari-ate. ويطلق على الصفات التى تتغير عشوائياً اسم المتغيرات العشوائية Random Variables وهى تلك التى تحدث بالصدفة مثل درجة الحرارة وشدة سقوط الأمطار. وتنقسم المتغيرات فى قيمها العددية إلى قسمين هما: المتغيرات المتصلة (المستمرة) Continuous Variables وهى المتغيرات التى يمكن أن تأخذ أى قيمة (أعدادصحيحة وكسور) على المقياس المستخدم. فمثلاً إذا إرتفعت درجة الحرارة يوم من ٢٠ درجة مئوية إلى ٣٠ درجة مئوية خلال الترمومتر الزئبقى، فمعنى ذلك أن الزئبق يكون قد مر بكل القيم الواقعة بين هاتين الدرجتين. كذلك الحال فى مقياس سرعة السيارة، فإذا زادت السرعة من ٣٠ كيلو متر/ ساعة إلى ٦٠ كيلو متر/ ساعة فإن المؤشر في المقياس يكون قد مر على كل القيم المحصورة بين هذين الرقمين. وبالمثل أيضاً الأطوال، وذلك لأن طول الشخص قد يكون ١٦٨ سنتيمتر أو ١٦٨,١ أو أى قيمة مهما كانت كسرية، وأصغر من المليمتر إذا كان المقياس يسمح بذلك. والنوع الآخر من المتغيرات يطلق عليه المتغيرات غير المتصلة أو الوثابة Discrete Variables وهى التى تختلف قيمها من مرحلة إلى أخرى بدون أن تكون منتظمة، كما أن قيمها لاتأخذ إلا أعداداً صحيحة Integras. فعدد الرحلات التى يقوم بها الأشخاص، وكمية الفيضان فى الأودية الصحراوية، وعدد الأنهار والبحار والقارات، وعدد السيارات المارة فى أحد الشوارع، وعدد الفصول فى المدارس وعدد الحجرات بالمنازل، وحجم الأسرات ... إلخ، كلها متغيرات وثابة (غير متصلة) نحصل عليها فى الغالب بالعدد، ولذا تعرف بيانات هذه المتغيرات أحياناً بالبيانات العددية Enumeration Data.

أنواع البيانات:

تمثل الطرق المختلفة التى تقاس بواسطتها البيانات أهمية خاصة لأساليب التحليل الإحصائى الكمى، إذ أن لكل أسلوب منها طريقة خاصة تقاس أو تجمع على أساسها البيانات. وبشكل عام هناك أربع طرق مختلفة تقاس بواسطتها أربع مجموعات رئيسية من البيانات هى: البيانات الإسمية، البيانات الترتيبية، بيانات الفترة، وبيانات النسبة.

(١) البيانات الإسمية (النوعية) Nominal Data:

تشتمل قياسات خصائص الظاهرة موضع الدراسة فى هذا النوع من البيانات على قياسات ثنائية أو ثلاثية. ولنضرب مثالاً على ذلك، فعند تسجيل الحالة التعليمية لمجموعة من الأشخاص، فإننا فى هذه الحالة نعطى للشخص الأمى الرقم (١) والشخص المتعلم الرقم (٢)، وإذا كانت الدراسة تتعلق بتسجيل نوع التربة: طينية أم رملية فى منطقة ما، نعطى التربة الطينية الرقم (١) والتربة الرملية الرقم (٢) وإذا لم تكن التربة طينية أو رملية تعطى الرقم (صفر).

ويطلق على البيانات الإسمية أحياناً إسم البيانات التصنيفية: لأنها تصنف المتغيرات على أساس خصائصها. ومن أمثلتها تصنيف التربة إلى التربة البنية، تربة البودزول، تربة اللاتريت ... إلخ، وتصنيف النباتات الطبيعية إلى غابات، حشائش وصحارى. وكلها أمثلة لا تحمل الصورة الكمية. وتعد طريقة قياس البيانات الإسمية من أضعف مستويات القياس لإستخدامها الأسماء أو الترقيم أو الرموز. وإستخدام الرموز فى مثل هذا النوع من البيانات ليس له أى علاقة بنوع الظاهرة أو درجة تركزها. فمثلاً فى وصف لون الحجر الرملى يمكن أن توضع بيانات الوصف كالآتى:

صفة اللون :	أحمر	رمادى	أصفر	أبيض
أو (بالرقم) :	١	٢	٣	٤
أو (بالرموز) :	؛	٪	:	*

(٢) البيانات الترتيبية Ordinal (Ranking) Data:

تعرف البيانات الترتيبية بالبيانات المرتبة فى فئات أو حسب خصائصها عن طريق إعطاء القيم الأصلية للمتغيرات رتباً أو أرقاماً تدريجية أو تنازلية. فمثلاً عند تصنيف المناطق المختلفة فى الوجه البحرى - مصر - حسب كمية الأمطار التى تسقط عليها، نجد أن هناك مناطق تسقط عليها أمطاراً كميتها أقل من ٢٥ ملليمتر، ٢٥ - ٧٥، ٧٥ - ١٢٥، أكثر من ١٧٥ ملليمتر، ففى هذه الحالة تعطى المنطقة التى تسقط عليها أقل من ٢٥ ملليمتر الرقم (١) والمنطقة التى

تسقط عليها ٢٥ – ٧٥ ملليمتر الرقم (٢)، وهكذا إلى أن نصل إلى المنطقة التى تسقط عليها أكثر من ١٧٥ ملليمتر فيعطى لها الرقم (٤). ونتبع نفس الأسلوب فى حالات أخرى كقياس الانحدارات وتقسيمها إلى انحدارات شديدة، انحدارات معتدلة، انحدارات طفيفة، وهكذا. ويستخدم هذا النوع من طرق قياس البيانات أيضاً مع الظاهرات التى تختلف طبيعة متغيراتها، والتى يمكن أن توضع فى شكل تتابع أو تسلسل تدريجى قائم على أساس العلاقة بينها. ويعد مقياس موس Mohs'scale لتعيين صلابة المعادن من أوضح الأمثلة للبيانات الترتيبية. وقد اختير لهذا المقياس عشرة معادن رتبت ترتيباً تصاعدياً حسب درجة صلابها النسبية. وتبدأ بالمعدن الأقل صلابة وهو التلك ودرجة صلابته النسبية (١) وتنتهى بالمعدن الأعظم صلابة وهو الماس ودرجة صلابته (١٠)، وعلى أساس هذا الترتيب فى الصلابة يمكن تقدير صلابة المعادن الأخرى تقديراً نسبياً. وعلى العموم، إذ علم الترتيب (التدرج) الحقيقى للفئات التى تجمع على أساسها البيانات، فإنه يمكن إعطاء أى أرقام تدريجية لتمثل الفئات. فمثلاً تتكون الرواسب الشاطئية من مجموعة من العناصر يمكن ترتيبها حسب حجم الحبيبات التى تتكون منها وإعطائها أرقاماً تدريجية كما يلى:

اسم العنصر	:	صلصال	غرين	رمل	حصى
أو	:	٤	٣	٢	١
الحجم	:	٨٥٦	٢٧٩	٥٢	١
أو	:	١	٢	٣	٤

(٣) بيانات الفترة Interval Data:

تعد بيانات الفترة أكثر أنواع البيانات الإحصائية شيوعاً واستخداماً فى أبحاث العلوم الأرضية ومن بينها الجغرافية. وبيانات الفترة تعكس القيم الأصلية للظاهرات فى شكل فئات لها أطوال (أو فترات) فيما بينها، كأعمار السكان فى فئات السن المختلفة، وبالمثل كميات الإنتاج الزراعى والصناعى، مساحات المزارع، كميات الأمطار؛ ودرجات الحرارة، فكلها بيانات يمكن أن تقاس على أساس فئات

محدودة. ويتميز هذا النوع من البيانات بأن النسبة بين أى فترتين مستقلين من وحدات القياس ونقطة صفر القياس لأحد مقاييس الظاهرة تكون مساوية لمثيلتها على مقياس آخر لنفس الظاهرة. فمثلاً، على الرغم من إختلاف صفر المقياس المئوى عن صفر المقياس الفهرنهيتى لدرجة الحرارة، فإن نسبة أى فترتين على أحد المقياسين تساوى نفس النسبة على المقياس الآخر.

درجات مئوية :	صفر	٢٠	٥٠	١٠٠
درجات فهرنهيتية :	٣٢	٦٨	١٢٢	٢١٢

$$\text{النسبة لفترتين (مئوية)} \quad \frac{١٠٠ - ٥٠}{٢٠ - \text{صفر}} = ٢,٥$$

$$\text{النسبة لفترتين (فهرنهيتية)} \quad \frac{٢١٢ - ١٢٢}{٦٨ - ٣٢} = ٢,٥$$

ومن المعروف أن هناك بعض الاختبارات الإحصائية التى لاتقبل إلا بيانات الفترة، بل أن معظم الأساليب الإحصائية مثل: تحليل التباين، معاملات الإرتباط، تحليل الإنحدار، تشترط أن تكون البيانات من نوع بيانات الفترة.

(٤) بيانات النسبة Ratio Data:

يمكن اعتبار بيانات النسبة من نوع بيانات الفترة التى تقاس على مقياس نسبى (Ratio Scale) والتى تكون فيها درجة صفر المقياس ذات قيمة أو درجة حقيقية (True Zero). وتعد قياسات كل من بيانات الكتلة Mass الأطوال، الأوزان، سرعة التيار النهرى، زوايا ميل وخط مضرب الطبقات الصخرية، من أمثلة هذه البيانات. فمثلاً إذا كان أحد الكتب يزن رطلاً وكتاب آخر يزن رطلين، فإن نسبة وزنيهما تكون $\frac{١}{٢}$ = ٥,٠. وإذا استخدمنا المقياس المترى فإن وزنيهما سيكون ٤٥٣ جرام و ٩٠٦ جرام على الترتيب. وفى هذه الحالة ستظل النسبة بينهما ٥,٠ ($\frac{٤٥٣}{٩٠٦} = \frac{١}{٢}$). أما بالنسبة لبيانات الفترة التى لايمكن

قياسها على أساس نسبى، فمن أمثلتها درجة الحرارة. فمثلاً إذا كانت درجة الحرارة عند وقت الظهيرة ليومين متتاليين هى ١٠° مئوية و ٢٠° مئوية، فإن النسبة بينهما هى $\frac{١٠}{٢٠}$ = ٠,٥ ولكن ستختلف هذه النسبة لنفس درجتى الحرارة بالمقياس الفهرنهيتى أى $\frac{٥٠}{٦٨}$ = ٠,٧٣٥ .

ومما تجدر الإشارة إليه أن طريقة القياس النسبى للبيانات تعتبر أكثر تنوعاً وقوة من أى طريقة قياس أخرى تجمع على أساسها البيانات الإحصائية، لإنها تحقق الحصول على أكبر كمية من المعلومات، كما أنها تسمح بتطبيق أكثر الأساليب الإحصائية دقة.

مما سبق يمكن القول أن بيانات الصفات المميزة Attributes للظاهرات تقاس بمقياس البيانات الأسمية والترتيبة وتأخذ قيما غير متصلة أو وثابة. أما المتغيرات Variables فتقاس بمقياس بيانات الفترة والنسبة وتأخذ قيماً مستمرة (Yale and Kendall, 1953).

الدقة والأخطاء فى البيانات Precision and Errors

يشير معنى الدقة Precision فى الإحصاء إلى التكرار فى عملية القياس، والتى ينتج عنها قيمة تقديرية تمثل القيمة المتوسطة لعدد مرات قياس المتغير موضع الدراسة. وكلما كانت طريقة القياس مضبوطة (دقيقة) كلما كانت المفردات الناتجة عن عملية القياس قريبة من بعضها أو تقترب جميعها من قيمة متوسطة Average Value. ونظراً لأن تكاليف وغرض عملية القياس غالباً ما تتحكم فى مقدار الدقة المطلوبة فإنه ليس بالضرورة أن تكون القياسات الدقيقة لمتغير ما صحيحة أو مضبوطة Accurate. ويعزى السبب فى ذلك إلى أن القياس الصحيح أو المضبوط يتوقف على تقدير قرب القيمة الدقيقة من القيمة الحقيقية، أو بمعنى آخر عدم تحيز القياس، الظاهرة المقاسة. وفى الدراسات الجغرافية تواجهنا مشكلة عدم معرفة القيم الحقيقة للظاهرة. ولكن فى بعض الحالات يمكن افتراض قيمة معيارية تقارن بها بقية القيم أو القياسات. فعلى سبيل المثال يمكن قياس انحدار جزء من الأرض

طوله ٥٠٠ متراً بدقة كبيرة إلى أقرب ملليمتر، ومع ذلك إذا كانت أداة القياس (الشريط) غير مرقمة ترقيماً صحيحاً، أى نتج عنها خطأ أو تحيز، فإننا نحصل على قياس دقيق ولكنه غير صحيح. وتبعاً لذلك يتعرض العمل الإحصائى إلى أنواع كثيرة من الأخطاء Errors أثناء تنفيذه، وسنشير هنا إلى نوعين رئيسيين من أنواع الأخطاء التى يتعرض لها قياس البيانات، والتى من شأنها التأثير على النتائج التى نحصل عليها من العينات وهما: أخطاء التحيز Bias Errors والأخطاء الإحتمالية Probability Errors.

وأخطاء التحيز هى الأخطاء الناجمة عن تدخل الباحث فى طريقة اختيار العينة. فالمعروف مثلاً أن العينة العشوائية تمثل بشكل كبير خصائص المجتمع الذى سحبت منه. فإذا اختيرت العينة بطريقة شخصية (أى غير عشوائية)، فإن ذلك يؤدى إلى زيادة الأخطاء المتوقعة. كذلك تنشأ هذه الأخطاء نتيجة لتحيز الباحث لوجهة نظر خاصة تجاه القرارات المتخذة. ويحدث خطأ التحيز عادة فى اتجاه واحد إما بالزيادة أو بالنقص. ويمكن أن تعزى أخطاء التحيز لعدة عوامل أهمها:

١- الاختيار المتعمد (غير العشوائى) للعينة.

٢- استبدال مفردات العينة بمفردات أخرى لعدم تمكن الباحث من الوصول لبعض المفردات الأساسية فى العينة.

٣- سوء التقدير وعدم توفر الدقة Precision، فقد لايوفق الباحث فى التفرقة بين ما هو سبب أو نتيجة، أو عدم توفر الدقة فى حصر وحساب المتغيرات المحددة لطبيعة الظاهرة، ووضع فروض غير سليمة.

أما الأخطاء الاحتمالية فهى الأخطاء الناجمة عن احتمالات عدم تماثل النتائج التى نحصل عليها مع خصائص المجتمع. فحتى عندما تؤخذ العينة بالأسلوب العشوائى فإنه تظل هناك احتمالات أخطاء فى مدى تمثيل العينة لخصائص المجتمع الذى أخذت منه. ومن أهم هذه الأخطاء ما يطلق عليه إحصائياً «خطأ

الصدفة» أو «الخطأ العشوائي». وهو الخطأ الذى ينشأ فى عملية اختيار العينة. فإذا فرضنا أننا أخذنا من مجتمع الطلبة فى أحد قاعات الدراسة أربعة أفراد كانت أعمارهم كالآتى: ٢١، ١٩، ١٧، ٢٢ نجد أن متوسط عمر الفرد فى المجموعة المختارة هو $\frac{٧٩}{٤}$ = ١٩,٧٥ سنة. ونفرض أننا رمزنا له بالرمز ل. فإذا أخذنا من بين هؤلاء الأربعة ثلاثة أفراد عشوائياً، نجد أن هناك طرقاً عددها = ٤ق٣ = $\frac{٤ل٣}{٣ !}$ أى $\frac{٤ \times ٣ \times ٢}{٣ \times ٢ \times ١}$ = $\frac{٢٤}{٦}$ = ٤ طرق. فلايمكن إذن سحب العينة إلا بإحدى الطرق الأربعة التى تكون المجموعات (١، ٢، ٣)، (١، ٢، ٤)، (١، ٣، ٤)، (٢، ٣، ٤). وإذا حسبنا متوسط عمر الفرد فى كل المجموعات الأربع الممكنة نجده عل يالترتيب ١٩، ٢٠,٦٦، ٢٠، ١٩,٣٣. فإذا رمزنا إلى كل متوسط من هذه المتوسطات بالرمز ع، نجد أن ع تختلف عن ل إما بالزيادة أو بالنقص كالآتى، وعلى الترتيب: – ٠,٧٥، + ٠,٩٢، + ٠,٢٥، – ٠,٤٢ ويلاحظ أن المجموع الجبرى لهذه الاختلافات يساوى صفراً، وهذا يعنى أن الاختلاف (ع – ل) يكون فى المتوسط لكل العينات الممكنة مساوياً للصفر.

ويطلق على اختلاف التقدير (ع)عن القيمة الحقيقة أو قيمة المجتمع بالخطأ العشوائى أو خطأ الصدفة لأن هذا الخطأ يرجع إلى عملية اختيار العينة كما ذكرنا. وعلى أية حال يتوقف الخطأ العشوائى على عاملين رئيسيين هما:

١– حجم العينة: فكلما كبرت العينة كلما قل الخطأ العشوائى، وزادت الثقة فى النتائج التى نحصل عليها.

٢– تباين مفردات المجتمع: فكلما زاد تباين مفردات المجتمع زاد احتمال حدوث الأخطاء العشوائية. وأياً كان الأمر، لو اتبع الأسلوب العشوائى فى اختيار العينة لأصبح من السهل والإمكان تقدير هذا النوع من الخطأ من العينة ذاتها. وسيتضح ذلك جلياً عند دراسة تقدير معالم المجتمع Population parameters فى الفصول القادمة.

الباب الأول
جمع البيانات وطرق إعدادها للتحليل الكمى

مقدمة

الفصل الأول: جمع البيانات

الفصل الثاني: تصنيف وجدولة البيانات

الفصل الثالث: العرض البياني للبيانات

الباب الأول
جمع البيانات وطرق إعدادها للتحليل الكمى

مقدمة:

يمر أسلوب أو خطة العمل الإحصائى أو ما يطلق عليه أحياناً «العمليات الإحصائية» بأربع مراحل رئيسية هى:

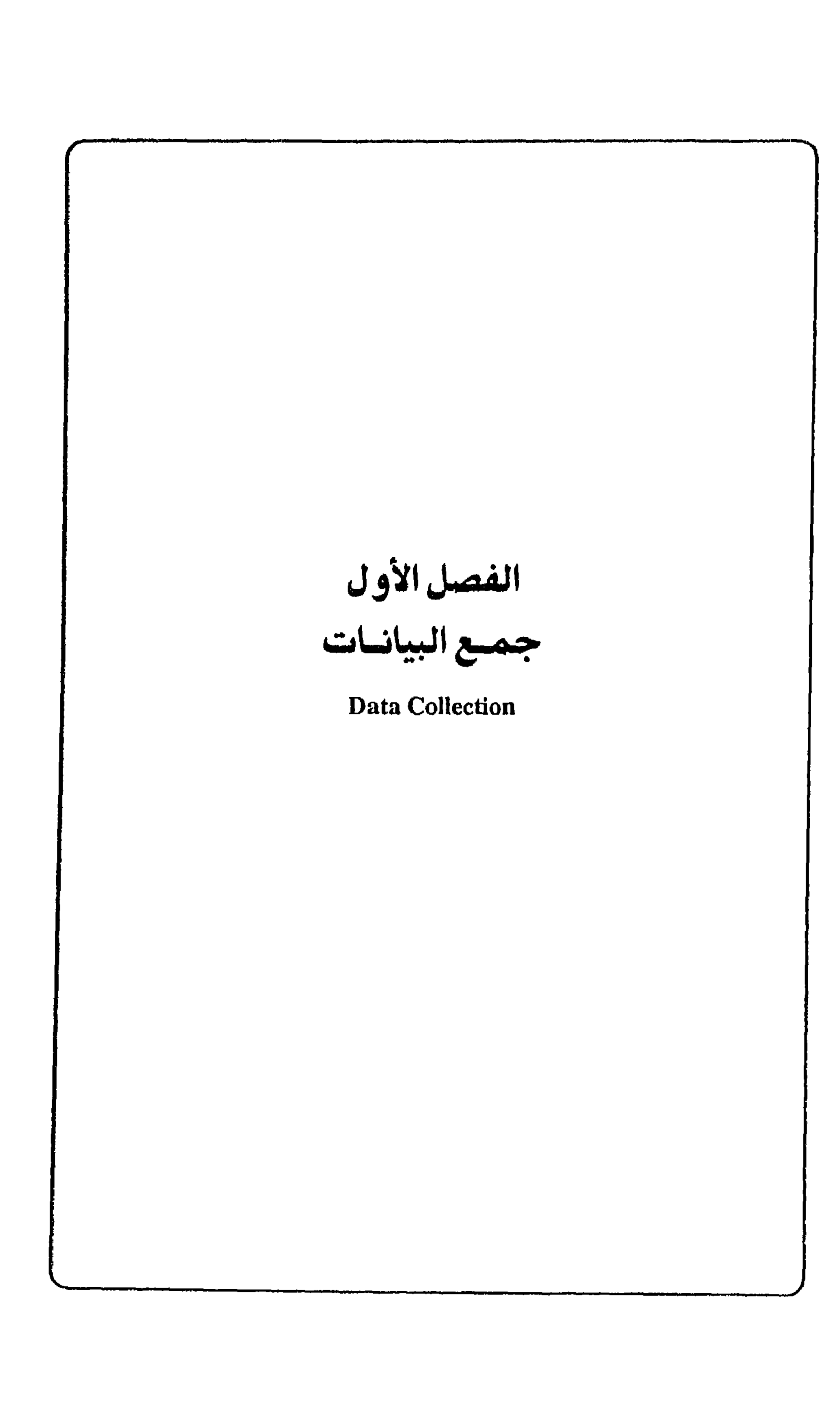
(١) مرحلة جمع البيانات Data Collection عن الظاهرة (الظواهر) موضع البحث من مصادرها المتنوعة سواء بواسطة المجهود الشخصى للباحث، أو عن طريق البيانات المنشورة.

(٢) مرحلة جدولة وعرض البيانات Data Tabulation and Presentation بما تتضمنه من طرق تفريغ المعلومات بأساليب وأشكال (جداول . رسوم أو أشكال بيانية ... إلخ) تعكس بشكل واضح وبسيط خصائص الظاهرة موضع الدراسة.

(٣) مرحلة تحليل البيانات Data Analysis بما تشمله من متغيرات يتأثر بعضها البعض، وعلاقات متداخلة مع بعضها البعض. ويتم التحليل بواسطة المقاييس الإحصائية والأساليب الكمية المتنوعة.

(٤) مرحلة تفسير البيانات Data Intepretation، وهى أوج المراحل الثلاث السابقة، كما أنها تهدف إلى معرفة العوامل التى تتحكم فى تنوع الظاهرة وتغير سلوكها.

وكل مرحلة من هذه المراحل تتضمن عدة طرق وأساليب في تنفيذها. وسوف نقوم بتوضيح وتفصيل المرحلتين: جمع البيانات، جدولة وعرض البيانات في فصول هذا الباب من الدراسة، أما بالنسبة للمرحلتين: تحليل البيانات وتفسير البيانات فسيتم شرحهما تباعاً في فصول الأبواب اللاحقة.



الفصل الأول

جمـع البيانـات

Data Collection



جمـــع البيانـــات

Data Collection

تعتبر مرحلة جمع البيانات والمعلومات والحقائق عن المتغيرات والظواهر موضع الدراسة من أسس العمل الإحصائى التى لها أهمية خاصة لايمكن إغفالها فى أى دراسة علمية منظمة. وقبل الشروع فى عملية جمع البيانات يجب أن يلم الباحث بعدة خطوات هامة وضرورية تمليها عليه طبيعة الدراسة يمكن أن نوجزها فيما يلى:

أ- تحديد المشكلة العلمية أو تعيين مجال الظاهرة المراد دراستها وبحثها.

ب- الإتفاق على وحدة القياس التى ستستعمل فى عملية جمع البيانات.

حـ- تعيين المتغيرات التى ستتناولها عملية القياس والحصر.

د- حصر المصادر التى يعتمد عليها فى الحصول على البيانات.

هـ- تحديد الأسلوب أو الطريقة التى تتبع فى جمع البيانات والمعلومات.

وسوف نركز مناقشتنا فى هذا الفصل حول الإطار العام لكيفية جمع البيانات من مصادرها المختلفة وما يتصف به كل مصدر من مزايا الاستخدام ومثالب ومشاكل التطبيق. وتجدر الإشارة هنا إلى أنه كلما كانت طريقة جمع البيانات سليمة، وكلما توفرت معلومات دقيقة عن مجموعة المتغيرات أو الظاهرة موضع الدراسة، كلما أدى ذلك إلى رفع درجة الثقة فى النتائج المستخلصة من التحليل الإحصائى، وبالتالى التوصل إلى قرارات سليمة غير متحيزة.

مصادر جمع البيانات Sources of Data

هناك مصدران أساسيان لجمع البيانات: الأول، يستمد منه الباحث المعلومات اللازمة لبحثه من بيانات تم جمعها وتجهيزها ونشرها بواسطة أجهزة متخصصة أما المصدر الثانى فيعتمد فيه الباحث على نفسه فى جمع وإعداد وتجهيز البيانات. ويعرف المصدر الأول بالمصدر غير المباشر، بينما يطلق على المصدر الثانى «المصدر المباشر» أو مصدر الميدان.

أولاً: المصدر غير المباشر فى جمع البيانات:

تتصف البيانات التى نحصل عليها من هذا المصدر بأنها بيانات غير أولية، تم تبويبها وتصنيفها من قبل بواسطة شخص آخر (غير الباحث) أو هيئة حكومية، ومن أمثلتها البيانات التى تتضمنها الدوريات والنشرات والكتب والتقارير والبحوث التى تصدرها وتنشرها الجهات والهيئات الحكومية ومراكز البحوث العلمية. ويلجأ الباحث إلى هذا المصدر فى الحصول على البيانات التى يحتاج إليها بحثه فى حالة وجود صعوبات (من حيث الوقت والتكاليف) تعترض عملية جمع البيانات من مصادرها الأولية. وعلى الرغم من سهولة وسرعة الحصول على البيانات من هذا المصدر، إلا أنه يعاب عليه صعوبة تحديد درجة الدقة أو الثقة فى البيانات، وعدم التأكد من سلامة الإعداد والتجهيز الإحصائى لها. وللتغلب على كل ذلك على الباحث أن لايتمادى فى الاعتماد على هذا المصدر فى حصوله على البيانات، وإذا كان مضطراً لذلك فيجب عليه الإعتماد على البيانات التى تصدرها أجهزة الإحصاء الرسمية فى الدولة، مثل الجهاز المركزى للتعبئة العامة والإحصاء بجمهورية مصر العربية.

ثانياً: المصدر المباشر فى جمع البيانات:

تتميز البيانات التى يتم الحصول عليها من هذا المصدر بأنها بيانات أولية يعتمد الباحث فى جمعها وتجهيزها للتحليل على نفسه. ويلجأ الباحث إلى هذا المصدر

فى حالة إذا ما كانت طبيعة الدراسة تملى عليه الحصول على بيانات غير منشورة، أو نتائج بحوث سابقة تتعلق بموضوع البحث، كما فى دراسة العلاقة بين العمليات البحرية Marine Proceases (الأمواج، التيارات ... إلخ) والظاهرات التى تتأثر بها على ساحل منطقة ما فى وقت معين. ومن مزايا المصدر المباشر فى الحصول على المعلومات أن درجة الدقة وحدود الثقة فى البيانات يمكن تحديدها عند تحليل البيانات كمياً، وهى فى الغالب ما تكون مرتفعة مما يساعد بالتالى على استخلاص نتائج موثوق فيها بدرجة كبيرة. إلا أن أهم المشاكل التى تواجه الاعتماد على المصدر المباشر هو الحاجة إلى الوقت والتكلفة المادية اللازمين لإنجاز مهمة الحصول على المعلومات. ونتيجة لذلك فإن الباحث يجد نفسه مضطراً إلى بذل قصارى جهده فى جمع البيانات التى يحتاج إليها بالطريقة المباشرة فى وقت قصير وبأقل تكلفة مادية ممكنة.

وعند جمع البيانات من مصادرها المباشرة فإن الباحث يعتمد على أحد الأسلوبين: أسلوب الحصر (المسح) الشامل لجميع مفردات المجتمع الأصلى، فإذا لم يتيسر له ذلك فإنه يضطر إلى اختيار عينة، وهذا ما يطلق عليه أسلوب المعاينة (العينات). ولكل من الأسلوبين جوانبه الإيجابية والسلبية التى نوضحها فيما يلى:

أولاً: أسلوب الحصر (المسح) الشامل:

يعرف أسلوب الحصر الشامل أحياناً بأسلوب العد الكامل (أو التعداد Census) حيث أن معظم التعدادات تتم من خلاله، مثل التعداد السكانى Population Census والتعداد الزراعى أو التجارى أو الصناعى التى يعتمد عليها فى استخراج بعض المقاييس والمؤشرات الاحصائية، والتى تكون أساساً فى عملية التخطيط القومى أو وضع إطار عام للأبعاد الفعلية لإمكانية الدولة فى مواجهة الأزمات الاقتصادية أو الاجتماعية وغيرهما. والأساس فى عملية جمع البيانات عن طريق الحصر الشامل هو إدخال كل مفردات المجتمع الاحصائى، دون استبعاد أى مفردة، فى البحث والاستقصاء. فمثلاً عند دراسة العمالة الصناعية فى محافظة ما يقوم الباحث بعمل

حصر شامل لجميع العمال حسب نوع كل صناعة، وكذلك عند دراسة التركيب المحصول للأحواض الزراعية فى أحد مراكز محافظة ما فإن الباحث يقوم بعمل حصر شامل لأنواع المحاصيل والمساحة التى تشغلها داخل كل حوض من الأحواض الزراعية. وبناء على ذلك فإن هذا الأسلوب يطبق عند دراسة المجتمعات الاحصائية مجهولة المعالم والتى تتطلب جمع بيانات شاملة عن كل مفردة من مفردات المجتمع حتى يمكن تحديد خصائصه ومعالمه بكل دقة وبدرجة عالية من الثقة.

ولأسلوب الحصر الشامل بعض المثالب والمشاكل عند استخدامه فى جمع البيانات فهو لايصلح للأبحاث التى يقترن استخلاص النتائج منها بوقت محدد، أو بمعنى آخر أن هذا الأسلوب لايتناسب مع الأبحاث التى يكون فيها لعنصرى الوقت والتكاليف المالية أهمية خاصة وأثر كبير على استخلاص النتائج. وعلاوة على ذلك يتعرض تنفيذ أسلوب المسح الشامل فى جمع البيانات لكثير من الأخطاء التى من أهمها خطأ تحيز الباحث، سواء كان تحيز متعمد أو غير متعمد، الذى ينجم عن أخذ كل مفردات المجتمع فى الدراسة حيث وجود احتمالات الخطأ فى العد أو احتمالات تجاهل بعض المفردات مما يؤثر على دقة النتائج. وللتخلص من خطأ هذا الأسلوب يمكن تقسيم المجتمع إلى مجموعات متجانسة لها خصائص متشابهة ومميزات مترادفة، ثم يجرى البحث وعملية الحصر على كل قسم على حدة مع مراعاة التنسيق فى الدراسة بين كل الأقسام. وأخيراً فإن هذا الأسلوب يتطلب فى اجرائه توفر جهاز فنى احصائى كبير واعتمادات مالية ضخمة ووقت متسع، مما يفسر أن معظم الدراسات والأبحاث التى تعتمد على هذا الأسلوب فى انجازها لايقوم بها سوى أجهزة الإحصاء فى الدول مثل الجهاز المركزى للتعبئة العامة والإحصاء بجمهورية مصر.

ثانياً: أسلوب المعاينة (العينات) Sampling

سبق أن عرفنا أن دراسة المجتمعات الاحصائية تعتمد أساساً على أخذ كل

مفردات المجتمع للتعرف على خصائص ومعالم هذا المجتمع. وبصفة عامة فإن معالم أى مجتمع (وهى مقادير ثابتة للمجتمع الواحد ولكنها تتغير من مجتمع إلى آخر) هى التى تعطى لهذا المجتمع صفاته دون غيره. ونظراً لوجود صعوبات كثيرة تحول دون دراسة جميع مفردات المجتمع بواسطة أسلوب الحصر الشامل، فإننا نجرى دراستنا على جزء صغير من هذا المجتمع أو ما يسمى بالعينة Sample، وذلك اختصاراً للوقت وتوفيراً للجهد والنفقات. واتباع دراسة العينات أو أسلوب المعاينة يرفع من مستوى العمل البحثى ويجعله أكثر دقة، وذلك لأن دراسة عدد قليل من المفردات أو الحالات يتيح للباحث فرصة جمع معلومات دقيقة وكثيرة عن كل مفردة أو حالة. وعلى العموم فإنه إذا ما وجدنا أنه من الضرورى إجراء معاينة فإن رائدنا الأساسى يكون دائماً هو الحصول على عينة تعطى نتائجاً ذات دقة معينة بأقل تكاليف ممكنة، أو التى تعطى أعلى دقة بتكاليف محدودة.

ويفضل استخدام أسلوب المعاينة عند دراسة خصائص ومعالم المجتمعات اللانهائية مثل الوحدات الإنتاجية لإنتاج بعض الآلات، كما يفضل كذلك فى الأبحاث العلمية التى تتطلب تصور عام أو رأى عام حول قضية أو مشكلة يرى دراستها فى مجالات العلوم الطبيعية أو الاجتماعية. وفى كل من هذه الحالات يجب أن تكون العينات ممثلة تماماً للمجتمع ولاتخضع للاختيار الشخصى، وذلك حتى يمكن الحصول، بواسطة تطبيق الأساليب الكمية والمقاييس الاحصائية، على نتائج يمكن تعميمها على المجتمع الأصلى المراد تحديد معالمه بدرجة عالية من الدقة والثقة. وتجدر الإشارة هنا إلى أنه عند دراسة العينات فإن المقاييس التى تحسب من توزيع العينة المختارة (مثل الوسط الحسابى والانحراف المعيارى – سيأتى ذكرهما فيما بعد بالتفصيل) يسمى كل منها «احصائية» وقيمة كل «احصائية» تختلف من عينة إلى أخرى. وللتفرقة بين المقاييس التى تحسبها من العينة وتلك التى نحصل عليها من دراسة جميع مفردات المجتمع بطريقة الحصر الشامل، تسمى

الأولى «بالاحصائيات» Sample statistics، بينما تعرف الثانية «بالمعالم» Population Parameters.

ويتوقف نجاح استخدام وتطبيق أسلوب المعاينة على عدة أمور هامة هي: تقدير حجم العينة، كيفية اختيار مفردات العينة من المجتمع، وتحديد نوع العينة. وفيما يلي مناقشة تفصيلية لكل منها على حدة.

(١) تقدير حجم العينة:

تتفق آراء كثير من الاحصائيين على أن حجم عينة البحث يتوقف على مجموعة من العوامل تنحصر في: الغرض من البحث، حجم المجتمع الأصلي، مدى تباين الظواهر المختلفة في قطاعات المجتمع. درجة الدقة المطلوبة في البحث، البيانات المتاحة التي يمكن استخدامها في تعميم النتائج، والامكانيات المادية. ونظراً لعدم وجود اتفاق بين الباحثين على وضع حد معين على أساس علمي – أو إحصائي – يحدد الحجم المناسب أو الأمثل للعينة لكي تمثل المجتمع الذي تسحب منه تمثيلاً جيداً، فإن تقدير حجم العينة – على مستوى معظم الدراسات والبحوث – يعتبر واحدة من المشكلات الخاصة بأسلوب المعاينة وتطبيق الأساليب الإحصائية. وفي مجال العمل الاحصائي يوجد اتجاهان عند تقدير حجم العينة.

الاتجاه الأول: يعتمد على الخبرة السابقة للباحث في هذا المجال، حيث أظهرت خلاصة الخبرات والتجارب أن حجم عينة في حدود ١٠٪ إلى ١٥٪ من حجم المجتمع الأصلي يبدو ملائماً في معظم الدراسات والبحوث. ويتميز هذا الاتجاه في تقدير حجم العينة بسهولته كما أنه يفيد بعض الباحثين قليلي الخبرة في مجال العمل الاحصائي.

الاتجاه الثاني: يرتبط أساساً بنظرية الاحتمال Theory of probability، مما يتطلب من الباحث الإلمام بقدر وافر من المعلومات الإحصائية والرياضية حتى يستطيع استخدام الأساليب الإحصائية في تقدير الحجم الأمثل للعينة. ويعتمد هذا

الاتجاه على تحديد العوامل (المتغيرات) التى يتوقف عليها حجم العينة واعتبارها دلائل رئيسية أو مؤشرات أساسية لهذا الغرض، وهو أمر يغفله الاتجاه الأول تماماً. كما يعتمد هذا الاتجاه على توفر بعض المعلومات عن حجم ومعالم المجتمع الأصلى عن طريق العينات التجريبية أو الاسترشادية Experimental or Pilot Sample. وتتمثل أهم العوامل والمتغيرات الرئيسية المحددة لحجم العينة فى: نسبة الخطأ المسموح به (أو درجة الدقة أو الثقة)، معامل التشتت (أو الإنحراف المعيارى) بين مفردات العينة أو المجتمع إن أمكن، والاختلاف النسبى بين المتوسط الحسابى للعينة ومتوسط المجتمع. وتوضع هذه المتغيرات فى شكل صيغة رياضية تختلف باختلاف حجم العينة الاسترشادية، كما تترجم على هيئة معادلة.خاصة فى حالة إذا كان حجم المجتمع الأصلى الذى ستحسب منه العينة معلوماً.

فإذا ما تصورنا أن أحد الجغرافيين بصدد تقدير حجم عينة من مجتمع كبير غير محدود المفردات فإنه يقوم بسحب عينة استرشادية من هذا المجتمع وحساب بعض المقاييس الاحصائية منها لتقدير بعض خصائص أو معالم المجتمع، والتى عن طريقها يمكن تقدير حجم العينة المطلوب. فإذا كان حجم العينة الاسترشادية ٣٠ مفردة أو أكثر فإن أهم العوامل المحددة لحجم العينة المطلوب تتمثل فى:

أ- الانحراف المعيارى بين مفردات العينة أو الخطأ المتوقع لمتوسط قيم مفردات العينة، ومنه يمكن تقدير الانحراف المعيارى للمجتمع، أو ما يعرف «بأحسن تقدير Best Estpmate للانحراف المعيارى بين مفردات المجتمع، ويرمز له بالرمز $\hat{ع}$ ويحسب على أساس:

$$\hat{ع} = عـ \times \sqrt{\frac{ن}{ن - ١}} \quad \text{أو} = \sqrt{\frac{مجـ (س - \bar{س})^{٢}}{ن - ١}} \quad \text{أو} = \sqrt{\frac{مجـ س^{٢} - \bar{س} ن}{ن - ١}}$$

حيث عـ هى الانحراف المعيارى للعينة، س هى قيمة مفردة من مفردات العينة، $\bar{س}$ هى المتوسط الحسابى للعينة، ن هى الحجم الفعلى للعينة.

ب- خطأ المعاينة أو الخطأ المعيارى Sampling Error or Standard Error للمتوسط بين مفردات العينة أو مفردات المجتمع إن أمكن. وهو عبارة عن الدقة المطلوبة للتقدير الإحصائى من بيانات العينة، إذ أن تحديد حجم العينة يعتمد على الدرجة التى عندها يتجه متوسط العينة إلى الاختلاف والتباين عن متوسط المجتمع(١). ويرمز للخطأ المعيارى بالرمز (خ. م) ويحسب على أساس:

$$(\text{خ.م}) = \sqrt{\frac{\hat{\text{ع}}}{\text{ن}}} \times \sqrt{\text{١} - \text{ف}} \quad \text{أو} = \sqrt{\frac{(\hat{\text{ع}})^{٢}}{\text{ن}}} \times \sqrt{\text{١} - \text{ف}}$$

$$\text{أو} = \sqrt{\frac{(\hat{\text{ع}})^{٢}}{\text{ن}} \times (\text{١} - \text{ف})}$$

حيث ف تمثل نسبة حجم العينة إلى حجم المجتمع الأصلى ن. (أى $\frac{\text{ن}}{\text{ن.}}$) وتسمى هذه النسبة «نسبة المعاينة» Sampling Fraction أو معامل التصحيح لقيمة الخطأ المعيارى للمجتمع الأصلى الذى يجب أن يكون أقل من خطأ المعاينة للمتوسط. وكلما كان زاد حجم العينة واقترب من حجم المجتمع الأصلى كلما اقتربت ف من الوحدة (الواحد الصحيح) وأصبحت قيمة معامل التصحيح صفراً، وبالتالى فإن قيمة الخطأ المعيارى تصير صفراً أيضاً.

حـ- القيمة المعيارية لاحتمال وقوع خطأ مسموح به ويرمز لها بالرمز (ز). ويمكن تحديد هذه القيمة من جدول التوزيع الاحتمالى الطبيعى إذا كان مستوى الثقة Confidence Level الذى تعمم به النتائج على المجتمع معلوماً.

وإذا أخذنا فى الإعتبار المتغيرات الثلاثة السابقة فإن حجم العينة يمكن أن يتحدد فى ضوء تحديد الفارق الممكن التسامح فيه بين نتيجة العينة، وما هو كائن

---

(١) سنتناول مناقشة الخطأ المعيارى Standard Error مناقشة تفصيلية عند دراسة أساليب المقارنة فيما بعد.

فعلاً فى المجتمع Tolerance (أى الخطأ المعيارى)، ومستوى الثقة التى تعمم بها النتائج على المجتمع. ويمكن وضع هذا التصور لحجم العينة بحسب أن:

$$\text{(خ.م)} = \frac{\hat{\text{ع}}}{\sqrt{\text{ن}}}$$

$$\hat{\text{ع}} = \text{خ.م} \times \sqrt{\text{ن}}$$

$$\sqrt{\text{ن}} = \frac{\hat{\text{ع}}}{\text{خ.م}}$$

$$\text{ن} = \left(\frac{\hat{\text{ع}}}{\text{خ.م}}\right)^{2} \quad \dots \quad \dots \quad \dots \quad (1-1)$$

فإذا استطعنا أن نضع تقديراً مبدئياً للخطأ المعيارى (خ.م) الذى نرغب أن ننتهى إليه، وإذا استطعنا أيضاً سحب عينة استرشادية كبيرة (من الثابت احصائياً أنه إذا بلغ حجم العينة ٣٠ مفردة أو أكثر فإنه يمكن أن يعطى حدوداً مرتفعة من الثقة لتقدير متوسط المجتمع وانحرافه المعيارى من متوسط العينة وانحرافها المعيارى، ويرجع ذلك إلى أنه كلما كان حجم العينة كبيراً كلما أدى ذلك إلى تكوين توزيع طبيعى للعينة يتركز حول المتوسط الحقيقى للمجتمع حتى لو كان توزيع المجتمع غير طبيعى)، فإنه بإستطاعتنا تقدير الحجم الأمثل بالاستعانة بالمعادلة السابقة (١-١).

مثال ١ : على أساس عينة استرشادية مكونة من ١٠٠٠٠ فدان قدر أن متوسط إنتاج الفدان من القمح على مستوى الجمهورية هو ٥ أرادب. وكان أحسن تقدير للانحراف المعيارى ($\hat{\text{ع}}$) لإنتاج القمح على المستوى القومى ٢ أردب، وأن الخطأ المعيارى للمتوسط هو ٠٢, أردب. فلو افترضنا أننا نريد تقدير المتوسط القومى لإنتاج الفدان من القمح لأقرب ٤/١ أردب عند مستوى احتمالى ٦٨٢, (أى بمستوى الثقة ٦٨,٢ ٪) فإن أقل حجم مطلوب للعينة فى هذه الحالة يكون:

$$\text{حجم العينة (ن)} = \left(\frac{\hat{\text{ع}}}{\text{خ.م}}\right)^{٢}$$

$$= \left(\frac{٢}{,٢٥}\right)^{٢} = (٨)^{٢} = ٦٤ \text{ فداناً}$$

وعلى ذلك فإن أى عينة مكونة من ٦٤ فدان تكون كافية لإعطاء تقدير للمتوسط القومى (أى متوسط المجتمع) بدقة ± ١/٤ أردب وبمستوى ثقة ٦٨,٢ ٪.

ولو افترضنا، مرة أخرى، أن درجة الدقة المطلوبة لتقدير المتوسط القومى لإنتاج الفدان من القمح هى نفس الدقة السابقة (أى إلى أقرب ١/٤ أردب) ولكن عند المستوى الاحتمالى ٠,٩٥٤ (أى بمستوى الثقة ٩٥,٤٪) الذى تكون عنده حدود الثقة عبارة عن ± خطأين معيارين للمتوسط (أى ± ٢ × خ.م). ومعنى ذلك أن ٢ × قيمة الخطأ المعيارى للمتوسط لابد أن تساوى الدقة المطلوبة لحساب المتوسط العام وهى ٠,٢٥ أردب. وبعبارة أخرى فإن الخطأ المعيارى للمتوسط، لدرجة ثقة ٩٥,٤ ٪، يساوى $\frac{٢}{٠,٢٥}$ أى ٠,١٢٥ أردب. وبتطبيق المعادلة (١ -١) فإن حجم العينة يكون فى هذه الحالة.

$$\text{حجم العينة (ن)} = \left(\frac{\hat{\text{ع}}}{\text{خ.م}}\right)^{٢}$$

$$= \left(\frac{٢}{٠,١٢٥}\right)^{٢} = (١٦)^{٢} = ٢٥٦ \text{ فداناً}$$

وبناء على ذلك فإننا لكى نحصل على تقدير للمتوسط القومى لإنتاج الفدان من القمح لأقرب ١/٤ أردب بمستوى ثقة ٩٥,٤ ٪ فإن حجم العينة اللازم لابد أن يكون ٢٥٦ فداناً.

وهناك صيغة أو معادلة أخرى لتحديد الحجم الأمثل للعينة تأخذ فى اعتبارها المتغيرات السابق ذكرها والتى تحسب من عينة استرشادية يبلغ حجمها ٣٠ مفردها أو أكثر، وهذه الصيغة هى:

حجم العينة =

$$\left(\frac{\text{الانحراف المعيارى للعينة} \times \text{القيمة المعيارية لاحتمال خطأ مسموح به بدرجة معينة}}{\text{الدقة المطلوبة للتقدير الإحصائى أو الخطأ المعيارى}}\right)^{٢}$$

أى أن:

$$ن = \left[\frac{عـ \times ز}{خ_م}\right]^{٢} \quad \ldots \quad \ldots \quad \ldots \quad (١ - ٢)$$

مثال ٢: إذا كان الانحراف المعيارى لعينة استرشادية مكونة من ٣٠ عاملاً لدراسة مستوى المعيشة لمجتمع عمالى هو ١٠ جنيهات شهرياً وأن الخطأ المعيارى المسموح به لتقدير المتوسط العام للدخل الشهرى هو ٢,٥ جنيهاً وذلك بمستوى ثقة ٩٥,٤ ٪ فإن الحجم الأمثل للعينة الذى يحقق الدقة المطلوبة يمكن تقديره بعد تحديد قيمة (ز) المعيارية من جدول التوزيع المعتدل المناظرة لمستوى الثقة ٩٥,٤ ٪ وهى فى هذه الحالة تساوى ٢ تقريباً. وعلى ذلك فإن:

$$\text{حجم العينة (ن)} = \left[\frac{١٠ \times ٢}{٢,٥}\right]^{٢} = (٨)^{٢} = ٦٤ \text{ عاملاً}$$

وبما أن حجم العينة الاسترشادية ٣٠ عاملاً فإننا نحتاج إلى عدد ٣٤ عاملاً آخر ليصبح الحجم الفعلى للعينة ٦٤ عاملاً والذى منه يمكن تقدير المتوسط العام لدخل المجتمع العمالى قيد البحث بالدقة المطلوبة (أو الخطأ المعيارى للمتوسط) المشار إليها).

ويمكن أيضاً تحديد حجم العينة التى تحقق الدقة المطلوبة أو الخطأ المسموح به فى حساب المتوسط العام للمجتمع من إحصائية عينة تجريبية صغيرة يقل عدد مفرداتها عن ٣٠ مفردة. وفى هذه الحالة تؤخذ العوامل السابقة المحددة لحجم العينة فى الإعتبار مع اختلاف واحد وهو أن القيمة المعيارية المناظرة لاحتمال وقوع

خطأ مسموح به بدرجة معينة فى جدول التوزيع المعتدل يجب أن تستبدل بقيمة معيارية أخرى من جدول توزيع «ستيو دنت – ت» Student-t Distribution مناظرة لعدد من المفردات يقل عن مفردات العينة التجريبية بمفردة واحدة وعند مستوى الدلالة أو الثقة المطلوب.

وبناء على ذلك فإن:

$$\text{حجم العينة} = \frac{\text{مربع (الانحراف المعيارى للعينة × قيمة ت المعيارية)}}{\text{مربع (الدقة المطلوبة أو الخطأ المعيارى)}}$$

أى أن:

$$\text{ن} = \frac{(\text{عـ} \times \text{ز})^{٢}}{(\text{خ}_{\text{م}})^{٢}} \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad (١ - ٣)$$

مثال ٣: فى دراسة اجتماعية عن مستوى المعيشة لعمال إحدى الصناعات التى يبلغ عدد مصانعها ٩٢ مصنعاً، يراد معرفة حجم عينة من المصانع يمكن منه تقدير متوسط الدخل السنوى لجميع العمال وذلك فى حدود ٥٠ جنيهاً زيادة أو نقص عن متوسط دخل عمال مصانع العينة بمستوى ثقة ٩٥٪ (أى بمستوى دلالة ٥٪)، فإننا فى هذه الحالة نأخذ عينة تجريبية مكونة من دخول عمال ٢٥ مصنعاً ونحسب منها متوسط الدخل والإنحراف المعيارى ولنفترض أنه كان ٣٣٢ جنيهاً و ١٦١,٥ جنيهاً على الترتيب. وبما أن عدد مفردات العينة التجريبية أقل من ٣٠ مفردة، فلابد إذن من تعيين قيمة ت من جدول توزيع «استيودنت –ت» المناظرة لعدد مفردات يقل عن مفردات العينة بمفردة واحدة (أى ٢٥ – ١) وعند نسبة الخطأ المسموح بها وهى ٥٪. وباستخدام هذه المؤشرات فإن قيمة «ت» تساوى ٢,٠٦. وبذلك فإننا يمكن أن نطبق المعادلة (١–٣) لنحصل على حجم العينة المطلوب كما يلى:

$$ن = \frac{(١٦١,٥ \times ٢,٠٦)^٢}{(٥٠)^٢} = ٤٤,٢$$

وبناء على ذلك فإن حجم عينة مكونة من عمال ٤٥ مصنعاً تكون كافية لإعطاء صورة صادقة عن الدخل السنوى لعمال جميع المصانع. ومن المعادلة السابقة (١-٣) يمكن أن تستنتج أنه كلما كبرت قيمة الانحراف المعيارى للعينة التجريبية كلما زاد حجم العينة المطلوبة، والعكس يحدث مع الدقة المطلوبة أو الخطأ المعيارى لحساب المتوسط العام، أى أنه كلما قلت (انخفضت) هذه الدقة أو زادت قيمة الخطأ المعيارى كلما قل حجم العينة. ففي المثال السابق إذا انخفضت الدقة في حساب المتوسط العام لتصل إلى ١٠٠ جنيهاً زيادة أو نقصان من متوسط العينة، فإن حجم العينة المطلوب سيكون ١٢ مصنعاً فقط، وهو حجم كاف أيضاً لإعطاء صورة عامة عن متوسط الدخل السنوى لعمال جميع المصانع بالدقة المطلوبة.

كما يمكن تحديد حجم العينة عن طريق تحديد النسبة المئوية لوجود الظاهرة موضع الدراسة فى العينة التجريبية وتعيين مستوى الثقة التى تعمم بها النتائج على المجتمع، وتحديد الفارق الممكن التسامح فيه بين نتيجة العينة وما هو كائن فعلاً فى المجتمع، وهذا يتطلب تحديد الخطأ المعيارى للعينة على النحو التالى:

$$\text{الخطأ المعيارى (خ.م)} = \sqrt{\frac{أ \times ب}{ن}}$$

حيث أ هى النسبة المئوية لوجود الظاهرة، ب هى النسبة المئوية لعدم وجود الظاهرة أى ١٠٠ مطروحاً منها نسبة وجود الظاهرة، ن هى حجم العينة. ويمكن نقل هذه المعادلة على النحو التالى:

$$ن = \frac{أ \times ب}{(خ.م)^٢} \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad (١ - ٤)$$

وفى كل الحالات إذا استطعنا أن نضع تقديراً مبدئياً لدقة القياس أو الخطأ المعيارى الذى نرغب أن ننتهى إليه، وإذا استطعنا كذلك تقدير نسبة وجود الظاهرة فى الحالات المدروسة، فإنه باستطاعتنا تقدير الحجم الأمثل للعينة بالمعادلة السابقة (١ - ٤).

مثال ٤ : من عينة استرشادية لدراسة مدى تأثير البرامج التلفزيونية على ثقافة سكان أحد الأقسام الإدارية بمحافظة ما وجد أن النسبة المئوية لحائزى الأجهزة التليفزيونية هى ٨٦٪ بدقة تصل إلى ± ١٠٪، والمطلوب تحديد الحجم الأمثل للعينة التى يمكن عن طريقها دراسة هذا التأثير بمستوى ثقة ٩٥,٤٪. فيما أن مستوى الثقة التى تعمم بها النتائج على المجتمع هو ٩٥,٤٪ فإن حدود هذه الثقة Confidence limits عبارة عن ± خطأين معيارين (أى ± ٣ - خ.م) وقيمتيهما لابد أن تساوى ١٠٪. ومعنى ذلك أن الخطأ المعيارى فى هذه الحالة هو ٥٪ (أى $\frac{١٠٪}{٢}$)، وأن حجم العينة المطلوب حسب المعادلة (١ - ٤) هو:

$$\text{حجم العينة (ن)} = \frac{\text{أ} \times \text{ب}}{(\text{خ.م})^{٢}}$$

$$= \frac{٨٦ \times ١٤}{(٥)^{٢}}$$

$$= \frac{١٢٠٤}{٢٥} = ٤٨,٢$$

وعلى ذلك فإن الحجم الأمثل للعينة والذى نأمل أن يحقق الدقة المطلوبة فى هذا المثال هو ٤٩ حائزاً لأجهزة التليفزيون فى المنطقة موضع الدراسة.

وتستخدم صيغة أخرى لتحديد الحجم المناسب للعينة اعتماداً على خصائص بيانات عينة استرشادية يمكن منها تعيين النسبة المئوية لوجود الظاهرة موضع الدراسة بالإضافة إلى تقدير الخطأ المعيارى وتحديد مستوى الثقة التى تعمم بها النتائج. هذه الصيغة هى:

$$\text{ن} = \text{أ} \times \text{ب} \times \left[\frac{\text{ز}}{\text{خ.م}}\right]^{٢} \quad ... \quad ... \quad (١ - ٥)$$

حيث ز هى القيمة المعيارية لاحتمال وقوع خطأ مسموح به عند مستوى احتمالى معين:

مثال ٥: من عينة تجريبية تتكون من ٣٠ ناخباً وجد أن ١٢ ناخباً سيقومون باعطاء أصواتهم لمرشح الحزب (أ)، فإذا أريد تقدير نسبة الناخبين الذين سيدلوا بأصواتهم لانتخاب مرشح هذا الحزب من جملة الناخبين بدقة (أو خطأ معيارى) ± ١٪ وبمستوى ثقة ٩٥,٤٪ فإن حجم العينة المطلوب لتحقيق ذلك يتحدد على أساس:

أ- نسبة الناخبين فى العينة $= \frac{١٢ \times ١٠٠}{٣٠} = ٤٠٪$

ب- القيمة المعيارية (ز) لمستوى الثقة ٩٥,٤٪ = ٢ تقريباً.

حـ- الخطأ المعيارى للتقدير = ١٪

وعلى ذلك فإن:

$$\text{حجم العينة (ن)} = ٤٠ \times ٦٠ \left(\frac{٢}{١}\right)^{٢}$$

$$= ٢٤٠٠ \times ٤ = ٩٦٠٠ \text{ مفردة أو ناخباً.}$$

وتجدر الإشارة هنا إلى أن معظم استطلاعات الرأى Opinion Polls فى النواحى السياسية تعتمد على عينات يصل حجم أى منها إلى ٢٠٠٠ مفردة تقريباً، حتى يكون الخطأ المعيارى المتسامح فيه بين نتيجة العينة وما هو كائن فعلاً فى المجتمع ± ٢٪ أو أكثر قليلاً، وذلك قبل أن يدخل فى الاعتبار عوامل أخرى مثل خطأ التحيز أو تغيير الرأى فى آخر دقيقة تجاه موضوع الاستطلاع (مثل ترشيح عضو أحد الأحزاب الأخرى) من جانب نسبة من المبحوثين أو الناخبين.

كما يمكن تحديد حجم العينة عن طريق معرفة حجم المجتمع الأصلى فقط،

وذلك عن طريق تحديد مجموعة من العوامل أو المحددات الرئيسية التى يمكن أن نجملها فيما يلى:

أ- حجم المجتمع الأصلى الذى ستسحب منه العينة ويرمز له بالرمز ن.

ب- معامل الاختلاف بين مفردات العينة ويرمز له بالرمز (م) ويحسب على أساس:

$$(م) = \frac{\text{الانحراف المعيارى}}{\text{المتوسط الحسابى}} \times ١٠٠$$

حـ- مربع متوسط معامل الاختلاف المتوسط بين مفردات العينة ويرمز له بالرمز (مَ س)^{٢}، ويحسب على أساس:

$$(\acute{م}\ س)^{٢} = \left[\frac{م^{٢}}{ن} \times (١ - ف)\right]$$

حيث ن هى حجم العينة، ف هى نسبة المعاينة Sampling Fraction أى نسبة حجم العينة إلى حجم المجتمع الأصلى.

د- الفارق النسبى بين المتوسط الحسابى للعينة ومتوسط المجتمع ويرمز له بالرمز (ق)، ويمكن حسابه كما يلى:

الفارق النسبى = الجذر التربيعى لمربع متوسط معامل الاختلاف بين مفردات العينة × القيمة المعيارية للدقة المطلوبة بدرجة معينة

أى أن:

ق = (مَ س × ز)

وبناء على العلاقات الرياضية المتبادلة بين هذه العوامل الأربعة وحجم العينة فإننا يمكن أن نحدد حجم العينة المطلوب فى ضوء المعادلة الآتية:

$$\text{حجم العينة} = \frac{\text{حجم المجتمع الأصلى} \times \text{مربع القيمة المعيارية} \times \text{مربع معامل الاختلاف}}{\text{حجم المجتمع الأصلى} \times \text{مربع الفارق النسبى} + \text{مربع القيمة المعيارية} \times \text{مربع معامل الاختلاف}}$$

أى أن:

$$\text{ن} = \frac{\text{ن}_{\cdot} \times (\text{ز}) \times (\text{م})^{٢}}{\text{ن}_{\cdot} \times (\text{ق})^{٢} + (\text{ز})^{٢} \times (\text{م})^{٢}} \quad \text{... (١ - ٦)}$$

مثال ٦ : يريد أحد الباحثين تحديد حجم عينة من مجتمع محتوى على ٢٠٠٠ مفردة وذلك فى ضوء الافتراضات الآتية التى يرى أنها ضرورية لتطبيق الطرق الإحصائية واستخلاص النتائج التى على أساسها تتخذ القرارات اللازمة.

أ- معامل الاختلاف بين مفردات العينة (م) فى حدود ٢٠ ٪.

ب- متوسط معامل الاختلاف بين مفردات العينة (م س) فى حدود ٢ ٪

حـ- نسبة الخطأ المسموح به لاتزيد عن ٥ ٪ أى أن تعميم النتائج بثقة قدرها ٩٥٪.

د- بما أن القيمة المعيارية المناظرة لاحتمال وقوع الخطأ المسموح به وهو ٥ ٪ فى جدول التوزيع المعتدل (الطبيعى) تساوى ٢ تقريباً فإن:

الفارق النسبى (ق) = ٠٢, × ٢ = ٠٤,

وعلى ذلك فإن حجم العينة المطلوب يمكن تحديده كما يلى:

$$\text{حجم العينة (ن)} = \frac{٢٠٠٠ \times (٢)^{٢} \times (٠,٢)^{٢}}{٢٠٠٠ \times (,٠٤)^{٢} + (٢)^{٢} \times (٠,٢)^{٢}}$$

$$= \frac{٣٢٠}{٣,٣٦} = ٩٥,٢٥$$

وعلى ذلك فإن حجم العينة المناسب والذى يحقق افتراضات الباحث هو ٩٦ مفردة تقريباً من مجتمع يحتوى على ٢٠٠٠ مفردة.

وهناك طريقة أخرى لتحديد حجم العينة يمكن استخدامها إذا كان حجم المجتمع الأصلى معروفاً، وذلك بعد تحديد الحجم التقريبى للعينة والذى يتطلب:

أ- تحديد الدقة المطلوبة أو الخطأ الذى يمكن التسامح بين نتيجة العينة وما هو كائن فعلاً فى المجتمع.

ب- تحديد مستوى الثقة التى تعمم بها النتائج على المجتمع.

حـ- اختيار النسبة المئوية لوجود الظواهر قيد البحث (ح) التى تحقق أكبر رقم إذا ما ضربت في النسبة المئوية المكملة (١٠٠ - ح).

وتتخذ معادلة هذه الطريقة الشكل الآتى:

$$\text{ن}_{١} = \left[\frac{\text{ح} \times (\text{١٠٠٪} - \text{ح})}{(\text{خ.م})^{٢}}\right] \times (\text{ز})^{٢} \qquad \text{... (١ - ٧)}$$

حيث ح هى نسبة وجود الظواهر قيد البحث وتمثل ٥٠٪، $\text{ن}_{١}$ هى الحجم التقريبى للعينة.

وبعد الحصول على الحجم التقريبى للعينة، يتعين تحديد الحجم الفعلى لها فى ضوء حجم المجتمع الأصلى (ى) وذلك باستخدام معادلة التصحيح وهى.

$$\text{الحجم الفعلي للعينة (ن)} = \frac{\text{ن}_{١}}{١ + \frac{\text{ن}_{١} - ١}{\text{ى}}} \qquad \text{... (١ - ٨)}$$

مثال ٧: لنفرض أن أحد الباحثين يريد تحديد حجم عينة من مجتمع يحتوى على ٤٠٠٠ مفردة بناء على بعض الافتراضات التى رآها ضرورية فى هذا الصدد. هذه الافتراضات هي:

أ- نسبة الخطأ المسموح به (الدقة) فى حدود ± ٥٪.

ب- مستوى الثقة التى تعمم بها النتائج لاتقل عن ٩٥٪.

حـ- نسبة وجود الظاهرة موضع البحث فى العينة ٥٠٪ ونسبة عدم وجودها ٥٠٪ أيضاً.

وباستخدام هذه الافتراضات، والتى يمكن أن تتحقق من تحديد حجم مناسب للعينة، نجد أنه عند مستوى الثقة ٩٥٪ تكون القيمة المعيارية المناظرة فى جدول التوزيع المعتدل (الطبيعى) هى ٢ تقريباً.

وعلى ذلك فإن حجم العينة المطلوب يتحقق بتطبيق المعادلتين (١–٧)، (١–٨) كما يلي:

$$\text{الحجم التقريبي للعينة} = \frac{٥٠ \times ٥٠}{(٥)^٢} \times (٢)^٢$$

$$= \frac{٢٥٠٠ \times ٤}{٢٥} = ٤٠٠ \text{ مفردة}$$

$$\text{الحجم الفعلي للعينة} = \frac{٤٠٠}{١ + \frac{٤٠٠ - ١}{٤٠٠٠}}$$

$$= \frac{٤٠٠}{١,٠٩٩} = ٣٦٣,٩٤ \text{ مفردة}$$

وعلى ذلك فإن حجم العينة المناسب والذى يمكن أن يحقق افتراضات الباحث هو ٣٦٤ مفردة من مجتمع يحتوى على ٤٠٠٠ مفردة أى بنسبة ٩٪ تقريباً من حجم المجتمع الأصلي.

أما إذا تجمعت لدينا بيانات خاصة عن معالم المجتمع الأصلي المتوسط الحسابي والانحراف المعيارى الذى ستسحب منه العينة دون أن نتمكن من معرفة حجمه، فإننا نستخدم طريقة المنحنى المعتدل (الطبيعي) للحصول على الحجم المناسب أو الأمثل للعينة. ويوضح ذلك المثال التطبيقى الآتى:

مثال ٨: إذا كان متوسط الإنتاج القومي للقمح فى عام ما هو ٧,٢ أردب للفدان والانحراف المعيارى لهذا المتوسط هو ١,٥ أردب، فما هو حجم العينة بالفدان التى ينبغي اختيارها من محافظة ما بشرط ألا يكون خطأ الصدفة أكثر من ٥٪ وأن يكون متوسط الإنتاج فى العينة ٧,٠ أردب؟

نظراً لأننا افترضنا أن نسبة خطأ الصدفة هى ٥٪، فإنه يمكن إيجاد القيمة المعيارية (ز) والتى تعرف بالمتغير المعيارى Standard Variable من الجداول

الإحصائية الخاصة بتحديد المساحة تحت المنحنى المعتدل (الطبيعى)[١] المناظرة لقيمة خطأ الصدفة. ثم تطبق المعادلة الآتية:

$$\text{المتغير المعيارى} = \frac{\text{الفرق بين متوسطي المجتمع والعينة} \times \sqrt{\text{حجم العينة}}}{\text{الانحراف المعيارى لمتوسط المجتمع}}$$

أى أن

$$(\text{ز}) = \frac{\text{م} - \bar{\text{س}} \times \sqrt{\text{ن}}}{\text{ع}} \qquad \dots \qquad \dots \qquad (\text{١} - \text{٩})$$

وعلى ذلك فإن:

$$\text{١,٦٤} = \frac{\text{٧,٢} - \text{٧,٠} \sqrt{\text{ن}}}{\text{١,٥}}$$

$$\text{,٢} \sqrt{\text{ن}} = \text{٢,٤٦}$$

$$\sqrt{\text{ن}} = \frac{\text{٢,٤٦}}{\text{,٢}} = \text{١٢,٣}$$

$$\text{ن} = \text{١٥١,٢٩} \ \text{فداناً}$$

بمعنى أنه ينبغى أن تسحب عينة حجمها ٥٢ فداناً من المحافظة قيد البحث لكى تحقق الدقة أو خطأ الصدفة والمتوسط المطلوب للعينة.

من كل ما سبق يمكن القول أن حجم العينة الذى نحصل عليه بإحدى المعادلات السابقة لايعتبر ملزماً لأن الافتراضات التى تقوم عليها هذه المعادلات غير ملزمة لأى دراسة، وكل ذلك ما هو إلا مجرد علامات تحدد أسلوب العمل فى هذا المجال فى حدود أقل خطأ ممكن وبطريقة موضوعية غير متحيزة.

(١) انظر جدول تحديد المساحة تحت المنحنى المعتدل (الطبيعى) ضمن ملاحق هذا الكتاب.

(٢) اختيار مفردات العينة:

بعد أن تعرفنا على الطرق المختلفة التى تحدد الحجم المناسب للعينة التى سيجرى عليها البحث والاستقصاء، فإننا الآن بصدد التعرف على طريقة (عملية) اختيار مفردات هذه العينة من بين مفردات المجتمع الأصلى، أو ما يعرف بأسلوب سحب العينة من المجتمع. وعملية اختيار مفردات العينة، كواحدة من المشكلات الخاصة بأسلوب المعاينة، تتوقف أساساً على حجم المجتمع الأصلى. فإذا كان حجم المجتمع صغيراً أى مشتملاً على عدد محدود Finite من المفردات، فإن المشكلة لاتكون مشكلة اختيار العينة من بين مفردات المجتمع، بل تكون مشكلة الحصول عدد كاف من المفردات لغرض البحث. فمثلاً إذا أراد الباحث أن يجرى دراسة على كبار المزارعين بإحدى القرى، كنموذج لنفس الفئة فى القطر كله، فقد يحدد هذه الفئة بأنها تشتمل على كل من يمتلك ١٠٠ فداناً أو أكثر من الأراضى الزراعية فى القرية. وفى هذه الحالة قد يكون عدد هؤلاء الملاك قليلاً لدرجة أن العينة تستنفذهم جميعاً. كما تكون عملية الاختيار من المجتمع الأصلى عملية مشترطة بشروط تحدد المفردات (عدد الملاك) التى تتكون منها العينة المطلوبة. وبالطبع كلما كثرت الشروط اللازمة للعينة كلما صعب الحصول عليها وكلما قل عدد المفردات التى يتم الاختيار من بينها. أما إذا كان حجم المجتمع الأصلى كبيراً جداً أى مشتملاً على عدد غير محدود من المفردات المستوفية لجميع الشروط اللازمة فى العينة فإنه من اللازم إجراء عملية اختيار مفردات العينة إما بواسطة الاختيار غير العشوائى (المعاينة العمدية) أو بواسطة الاختيار العشوائي. وقبل أن نوضح كيفية الاختيار فى كل من الطريقتين، فإنه يجدر بنا أن نذكر الشروط التي ينبغي توافرها فى العينة حتى نستعيض بها عن المجتمع الأصلي الكبير.

سبق أن قلنا أنه يجب أن تكون العينة ممثلة للمجتمع أو بمعنى آخر يجب مراعاة أن تمثل العينة جميع مفردات المجتمع، وألا تكون متحيزة biased لجزء أو

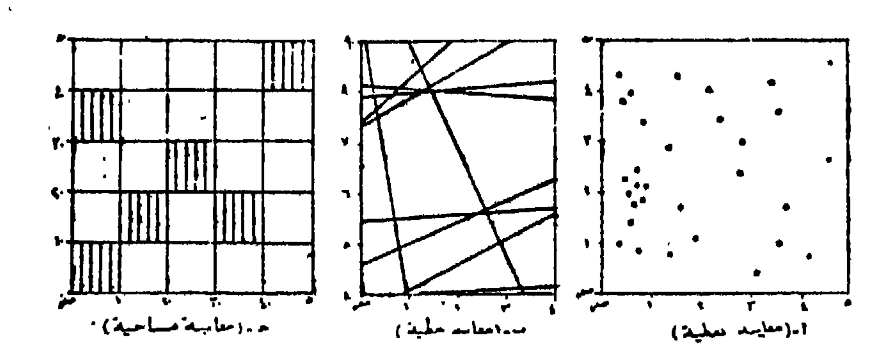
أجزاء من المجتمع الأصلي لأنه يتوقف على العينة المنتقاة كل قياس أو نتيجة يخرج بها الباحث. فمثلاً إذا أردنا سحب عينة لتقدير متوسط الدخل من الإنتاج الزراعى لملاك الأراضى الزراعية على مستوى أحد المراكز، فإذا أخذت عينة من فئة الملاك الذين يملكون ١٠٠ فداناً أو أكثر فإن العينة تكون غير ممثلة لمجتمع الملاك حيث أن هذه الفئة تمثل نسبة صغيراً جداً من جميع الملاك، وبالتالى لابد أن تحتوى العينة على ملاك من جميع فئات الملكية. وبناء على ذلك يجب أن تتوفر فى العينة الممثلة Representative Sample مجموعة من الشروط يمكن تلخيصها فى شرطين أساسيين هما:

١- أن تكون مفردات العينة ممثلة للمجتمع الذى يجرى عليه البحث تمثيلاً صحيحاً، وليست ممثلة لمجتمع آخر. بمعنى أنه إذا تكررت نفس النتائج على عينات أخرى من نفس المجتمع كانت العينة التى يجرى عليها البحث عينة ممثلة للمجتمع الأصلى أصدق تمثيل. وبذلك يمكن أن تكون خصائص مفردات العينة (احصائيات العينة) متقاربة أو متشابهة مع خصائص المجتمع (معالم المجتمع) الذى تنتمى إليه.

٢- ألا تكون المفردات المختارة ممثلة لجزء (قطاع) من أجزاء المجتمع الأصلى بل يجب أن تمثل جميع أجزاء المجتمع، وهذا يتطلب تكوين إطار المعاينة الذى تؤخذ منه العينة.

وإطار المعاينة Sampling Frame هو المصدر الذى تؤخذ منه العينة. أو بعبارة أخرى هو حصر شامل (القائمة أو الدليل) لجميع مفردات أو وحدات المجتمع الأصلى المراد دراسته. مثال ذلك قائمة بأسماء العمال فى أحد المصانع، أو مختلف أنواع الرواسب التى توجد على الشاطئ، أو مواقع المحلات العمرانية الريفية على خريطة إحدى الدول. وعند اختيار العينة من المجتمع المحدودة يقسم المجتمع الأصلى للظاهرة قيد البحث إلى عدة أقسام تسمى وحدات المعاينة Sampling Units (شخص، أسرة، قرية) ويكون إطار العينة حينئذ عبارة عن القائمة أو مجموعة

القوائم التى تتضمن الوحدات التى يتألف منها المجتمع. ويشترط فى إطار المعاينة أن يكون شاملاً لجميع مفردات المجتمع التى يمكن الوصول إليها بسهولة وذلك حتى يكون اختيار العينة سليماً. كما يشترط أن يكون إطار المعاينة متجدداً حتى تعطى المفردات أو الوحدات التى تستجد على الإطار القديم نفس الفرصة فى الظهور فى العينة. وكما ذكرنا فإنه فى المجتمعات غير المحدودة Infinit يستحيل إجراء حصر شامل لكل مفردات المجتمع فى الوقت المتاح للدراسة، ويكتفى فى هذه الحالة بدراسة عينة بدون تكوين إطار للمعاينة. ويلاحظ على إطار المعاينة فى مجال الدراسة الجغرافية أنه إما أن يكون إطاراً مكانياً Spatial أو غير مكانياً Non - Spa-tial. ويقصد بإطار المعاينة المكانى بأنه الإطار الذى يكون فيه المكان Location هو الوحدة الرئيسية، كما أنه الأساس فى اختيار العينات التى تمثل التغيرات (الاختلافات) المكانية Locational Variability التى يتميز بها مجتمع الأماكن لمنطقة ما تمثيلاً صحيحاً. فمثلاً إذا كنا بصدد معاينة خريطة بهدف تحديد مواقع الأماكن التى يشغلها نوعاً معيناً من النشاط البشرى على هذه الخريطة، فإننا يجب أن نتأكد من تمثيل كل أجزاء الخريطة تمثيلاً صحيحاً. ويتم ذلك باختيار أحد أنواع المعاينة الآتية (شكل رقم ١ – ١).



شكل رقم (١)

أنواع إطار المعاينة الجغرافية

أ- المعاينة النقطية Point - Sampling أى معاينة نقط مواقع الأماكن على خريطة المنطقة.

ب- المعاينة الخطية Line - Sampling أى بأخذ عينة من قطاعات عرضية مختلفة من الخريطة.

حـ- المعاينة المساحية Area - Sampling أى بأخذ عينة تمثل مساحة مجموعة من المربعات التى تغطى مساحة خريطة المنطقة قيد البحث.

وعلى ذلك يكون إطار المعاينة عبارة عن جميع مفردات المجتمع لكل شكل من أشكال المعاينة الثلاثة، أو بمعنى آخر جميع النقط المحددة لمواقع الأماكن، أو القطاعات العرضية، أو المربعات التى تغطى مساحة الخريطة. وعلى الرغم من أن طبيعة عمل الجغرافى عند جمعه للبيانات ترتبط بإطار المعاينة المكانى، إلا أنه فى بعض الأحيان ولظروف خاصة نجده يهتم بتحديد إطار معاينة غير مكانى ليلائم دراسته. فمثلاً إذا كان بصدد اختيار عينة من أسر أحد الأقسام الإدارية فى مدينة ما وذلك لتقدير متوسط الدخل، فإنه يتحتم عليه اختيار عينة من إطار (أو قائمة) يحتوى على جميع أسر هذا القسم الإدارى بالمدينة. ولايجوز له فى هذه الحالة أن يختار العينة من دليل التليفون مثلاً إذا أنه من المعروف أن مثل هذا الدليل لايتضمن جميع أسر القسم الإدارى قيد البحث.

الاختيار غير العشوائى (العمدى):

يلجأ الباحث أحياناً إلى اختيار مفردات عينة بطريقة غير عشوائية أو متعمدة، فمثلاً قد يختار الباحث عينة يرى أنها تمثل المجتمع بالنسبة إلى صفة أو خاصية ما دون غيرها. وبعبارة أخرى يكون الأساس فى الاختيار هو الباحث الذى يحدد بنفسه المفردات الداخلة فى العينة متحيزاً لتفكيره ومتعمداً فى تحديد المفردات. فمثلاً إذا أراد باحث دراسة مستوى المعيشة فى الريف المصرى فقد يعتقد أن قرية معينة فى نظرة تمثل مستوى المعيشة فى كل الريف المصرى، وفى هذه الحالة إذا اختار هذه

القرية كعينة أو أساس للدراسة فإن وجهة نظر الباحث تكون غير صائبة ولن يحالفها التوفيق في استخلاص النتائج لأنه باختياره لهذه القرية يكون معتمداً أو متحيزاً في الاختيار لغرض الدراسة ويمكن القول أن الباحث في تعمده في اختيار مفردات العينة إنما يتخلى عن فكرة العشوائية في اختيار مفردات العينة من مفردات المجتمع.

على أننا يجب ألا نقلل من أهمية الاختيار العمدى فربما يكون هو أفضل الطرق عند الاختيار في حالة إذا ما كان المطلوب اختيار عينة صغيرة لمجتمع كبير. فإذا كان المطلوب اختيار قرية واحدة لتمثيل القطر المصرى كله فإنه يمكن اعتبار الاختيار العمدى هو أفضل الطرق رغم ما فيه من تحيز، وذلك لأن اختيار قرية واحدة بطريقة عشوائية قد يؤدى إلى خطأ كبير.

الاختيار العشوائي:

على الرغم من سهولة اختيار أو سحب عينة غير عشوائية من المجتمع كله إلا أن ذلك له أضراره البالغة على دقة النتائج تبعاً لوجود تحيز من قبل الباحث في اختيار مفردات العينة ولعدم توفر عنصر العشوائية في الاختيار. ولضمان الحصول على المعاينة غير المتحيزة التي تعطينا تقديرات واستنتاجات يمكن تعميمها على جميع مفردات المجتمع الأصلى بأعلى دقة لتكاليف محدودة، لابد أن تختار العينة بحيث تكون ممثلة تمثيلاً صادقاً للمجتمع ويكون ذلك على أساس نظرية الاحتمالات، أى بواسطة سحب وحدات بالتتابع لكل منها احتمال معروف في الاختيار في السحبات المختلفة. ولكى تكون مفردات العينة ممثلة لكل مفردات المجتمع الأصلى بأقل أخطاء ممكنة، فلابد من استخدام أسلوب العشوائية غير المتحيز في الاختيار.

والأساس في أسلوب الاختيار العشوائي للعينة هو أن جميع مفردات المجتمع موضع الدراسة لها نفس الفرصة في الاختيار (Alder and Rossler, 1964) وهذا يعنى عدمالإهتمام ببعض المفردات دون الأخرى وإتاحة الفرص المتكافئة أمام كل مفردة لتكون ضمن العينة. والمعنى الرياضى في الاختيار العشوائى هو أن يكون

احتمال ظهور أى مفردة من مفردات المجتمع الأصلى فى العينة متساوى، ويساوى

$\frac{1}{\text{حجم المجتمع}}$، ويكون احتمال الحصول على مفردة من مفردات المجتمع فى العينة

يساوى $\frac{\text{حجم العينة}}{\text{حجم المجتمع}}$ $\left[\text{أى } \frac{ن}{ن}\right]$. وبذلك فإن الشرط الاحصائى الأساسى لاختيار مفردات العينة من بين مفردات المجتمع هو عدم التحيز فى الاختيار حتى نضمن - إلى درجة ما - تمثيل العينة للمجتمع الذى نريد معاينته تمثيلاً صادقاً مع تقليل الأخطاء الأخرى للمعاينة.

وتتم عملية اختيار مفردات العينة بالأسلوب العشوائى بإحدى الطرق الآتية:

أ- طريقة السحب العشوائى (القرعة).

عند اتباع هذه الطريقة تعطى مفردات المجتمع الأصلى أرقاماً مسلسلة تكتب على بطاقات متشابهة، وبعد أن تخلط خلطاً جيداً - يكفى لا ضاعة أى أثر للترتيب المتعمد - يسحب عشوائياً منها عدد عدد من البطاقات يساوى عدد مفردات العينة المطلوبة. وتلائم هذه الطريقة سحب العينات من المجتمعات الصغيرة الحجم حيث أنها لاتحتاج إلى مجهود كبير أو وقت طويل فى عملية ترقيم مفردات المجتمع الأصلى وسحب العينة منه.

ب- طريقة الجداول العشوائية:

يصعب اتباع طريقة السحب العشوائى فى الاختيار فى حالة المجتمعات الكبيرة الحجم حيث تحتاج عملية الترقيم إلى مجهود كبير كما أنها تكون مضيعة للوقت. ولذلك يفضل الرجوع إلى جداول خاصة تعرف باسم جداول الأرقام العشوائية Tables of Random Numbers. وهى جداول أرقامها مختارة بطريقة عشوائية (أى أنها أرقام لاترتبط ببعضها بأى أسلوب رياضى، فهى لاتكون بينها متتالية عددية أو هندسية) وموضوعة فى شكل أعمدة تتناسب مع حجم مجتمع أحصائى يتكون من أى عدد من المفردات كما يبدو من الجدول التالى (جدول رقم ١ - ١).

جدول رقم (١ -١) جدول الأرقام العشوائية

٢٠	١٧	٤٢	٢٨	٢٣	١٧	٥٩	٦٦	٣٨	٦١	٠٢	١٠	٨٦	١٠	٥١	٥٥	٩٢	٥٢
٧٤	٤٩	٠٤	٤٩	٠٣	٠٤	١٠	٣٣	٥٣	٧٠	١١	٥٤	٤٨	٦٣	٩٤	٦٠	٩٤	٤٩
٩٤	٧٠	٤٩	٣١	٣٨	٦٧	٢٣	٤٢	٢٩	٦٥	٤٠	٨٨	٧٨	٧١	٣٧	١٨	٤٨	٦٤
٢٢	١٥	٧٨	١٥	٦٩	٨٤	٣٢	٥٢	٣٢	٥٤	١٥	١٢	٥٤	٠٢	٠١	٣٧	٣٨	٣٧
٩٣	٢٩	١٢	١٨	٢٧	٣٠	٣٠	٥٥	٩١	٨٧	٥٠	٥٧	٥٨	٥١	٤٩	٣٦	١٢	٥٣
٤٥	٠٤	٧٧	٩٧	٣٦	١٤	٩٩	٤٥	٥٢	٩٥	٦٩	٨٥	٠٣	٨٣	٥١	٨٧	٨٥	٥٦
٤٤	٩١	٩٩	٤٩	٨٩	٣٩	٩٤	٦٠	٤٨	٤٩	٠٦	٧٧	٦٤	٧٣	٥٩	٢٦	٠٨	٥١
١٦	٢٣	٩٩	٠٢	١٩	٩٦	٤٧	٥٩	٨٩	٦٥	٢٧	٨٤	٣٠	٩٢	٦٣	٣٧	٢٦	٢٤
٠٤	٥٠	٦٥	٠٤	٦٥	٦٥	٨٢	٤٢	٧٠	٥١	٥٥	٠٤	٦١	٤٧	٨٨	٨٣	٩٩	٣٤
٣٢	٧٠	١٧	٧٢	٠٣	٦١	٦٦	٢٦	٢٤	٧١	٢٢	٧٧	٨٨	٣٣	١٧	٧٨	٠٨	٩٢
٠٣	٦٤	٥٩	٠٧	٤٢	٩٥	٨١	٣٩	٠٦	٤١	٢٠	٨١	٩٢	٣٤	٥١	٩٠	٣٩	٠٨
٦٣	٤٩	٠٠	٩٠	٦٧	٨٦	٩٢	٤٨	٣١	٨٣	١٩	٠٧	٦٧	٦٨	٤٩	٠٣	٢٧	٤٧
٦١	٠٠	٩٥	٨٦	٩٨	٣٦	١٤	٠٣	٤٨	٨٨	٥١	٠٧	٣٣	٤٠	٠٦	٨٦	٣٣	٧٦
٨٩	٠٣	٩٠	٤٩	٢٨	٧٤	٢١	٠٤	٠٩	٩٦	٦٠	٤٥	٢٢	٠٣	٥٢	٨٠	٠١	٧٩
٠١	٧٢	٣٣	٨٥	٥٢	٤٠	٦٠	٠٧	٠٦	٧١	٨٩	٢٧	١٤	٢٩	٥٥	٢٤	٨٥	٧٩
٣٧	٥٦	٤٩	٧٩	٣٤	٣٤	٣٢	٢٢	٦٠	٥٣	٩١	١٧	٣٣	٢٦	٤٤	٧٠	٩٣	١٤
٤٩	٠٥	٧٤	٤٨	١٠	١٥	٣٥	٢٥	٢٤	٢٨	٢٠	٢٢	٣٥	٦٦	٦٦	٣٤	٢٦	٣٥
٤٩	٧٤	٣٧	٢١	٩٧	٢٦	٣٣	٩٤	٤٢	٢٣	٠١	٢٨	٥٩	٥٨	٩٢	٦٩	٠٣	٦٦
٣٠	٢٦	٢٢	٤٣	٨٨	٠٨	١٩	٨٥	٠٨	١٢	٩٧	٦٥	٦٢	٦٣	٥٦	٠٧	٩٧	٨٥
٤٨	٨٧	٧٧	٩٦	٤٣	٣٩	٧٦	٩٣	٠٨	٧٩	٢٢	١٨	٥٤	٥٥	٩٣	٧٥	٩٧	٣٦
٠٨	٧٢	٨٧	٤٦	٧١	٧٣	٠٠	١١	٢٧	٧	٠٥	٢٠	٣٠	٨٥	٢٢	٢١	٠٤	٦٧
٩٥	٩٧	٩٨	٦٢	١٧	٢٧	٣١	٤٢	٦٤	٧١	٤٦	٢٢	٣٢	٧٥	١٩	٣٢	٢٠	٩٩
٣٧	٩٩	٥٧	٣١	٧٠	٤٠	٤٦	٥٥	٤٦	١٢	٢٤	٣٢	٣٦	٧٤	٦٩	٢٠	٧٢	١٠
٠٥	٧٩	١٨	٣٧	٨٥	٣٣	٧٥	١٨	٨٨	٧١	٢٣	٤٤	٥٤	٢٨	٠٠	٤٨	٩٦	٢٣
٥٥	٨٥	٦٣	٤٢	٠٠	٧٩	٩١	٢٢	٢٩	٠١	٤١	٣٩	٥١	٤٠	٣٦	٦٥	٢٦	١١

١٩	٠٨	٨٦	٦٤	٦٩	٠٥	٢٤	٤٩	٨٥	٣	٧٢	٢٤	٣٦	٦٨	٢٥	٩٦	٢٨	٦٧
٨٧	٠٦	٥٧	٨٩	٦٠	٤٥	٨٤	٧٣	٤٣	٨٦	٨٢	٥١	٥٩	٣٢	٧٨	٩٤	٨٦	٨٥
٦	٧٥	٤٧	١٨	١١	٣٧	٢٥	٥٨	١٣	٦٣	٤٤	٧٨	٨٨	٠٥	٠٩	٨٠	١٠	٤٠
١٩	٣٠	٢٠	٦٦	٤٧	٢٣	٤٤	٨٧	٨٩	٢٦	٨	٧٧	٨٠	٩٠	٤٨	٨٩	٥٥	٩٤
٠٩	١٤	٣٧	٥٦	١٥	٩٤	٠٣	٢٩	١٩	٦٢	٦٢	٣٣	٢٠	٢٣	٧٧	٧٧	٦٣	١١
٦١	٠٣	٣٥	٢٤	٠٤	٢٦	٤٠	٦٨	٦٢	٩٤	٥٧	٣٨	٥٥	٥٤	٠٤	٢٦	٠٠	٦٤
٦٩	٧٠	٩٨	٤٥	٢١	٣٠	٤٦	٨٨	٦٠	١٠	٥٨	٦٠	٤١	٧٨	٢٣	١٣	٩٤	٥٠
٨٥	١٧	١٩	٣٧	٩٧	٤٨	٨٥	٧٥	٥٩	٣٤	٠٥	٤٥	١٣	٤٤	٩٦	٤٧	٩٨	٦٦
٧٠	٠٨	٣٨	٧٢	٦٦	٥٧	٦٢	٧٩	٩٠	٦٠	٩١	٩٠	٧٧	٦٠	٨٣	٤٢	٩١	٦٦
٤٣	٨٥	٨٤	٥٨	٤٢	٩٠	٤٥	٣٩	٢٩	٢١	٤٠	١٩	٠٧	٠٢	١٨	١٢	٥٨	٣٣
٢٥	٣٧	٣٠	٠٥	٢٤	٩٨	٠٤	٠٢	٩٠	٥٠	٠٥	٧٣	٤٠	٧١	١٦	٤٠	٤٩	٥٢
٤١	٨٤	٧٦	٨٩	٣٧	٤٠	٥٨	٤٥	٩٢	٩٦	٤٧	٧٩	٣٠	٧٨	٩٩	٩٣	٩٨	٧٤
٤٤	٩٣	٦٩	٠٥	٢١	٢٣	٨٨	٦٩	٤٥	٥٤	٥٧	٦٩	٨٨	٠١	٣٠	٥٤	٢٦	٥٠
٣٤	٥٦	٥٩	٣٩	٧٣	٢٨	٣٥	١٩	٣٤	٢٨	٨٤	٩٦	٧٩	٣٣	٨٩	٦١	٤٦	٤٩
٩٠	٦٧	٧٦	٨٦	٩٦	٢٥	٠٠	٣٦	٠٣	٦٢	٨٨	٧٣	٣٩	٧٨	٤٤	١٣	٦٥	١٩
٣٨	٣٣	٨٦	٥٥	٢٨	٢٨	٠٠	١٤	٦١	٧٢	٤	٨١	٥٩	٨٧	٦٧	٤٧	١٧	٦٤
٦٠	٤٣	٠٧	٤٨	٨٩	١١	٨٨	٠١	٩٥	٥٧	٥٦	٥٦	٩٧	٦٨	٣٧	٩٧	٤٣	١٨
٠١	٣٧	٧٥	٤٤	٨٨	٦٦	٨١	٦٦	٥٣	١٥	٦١	٩٦	٠٩	٦١	٨٧	٦٠	٥٨	٦٥
٥٣	٣٤	٧٨	٦٤	٩٣	٤٥	١٥	٤٣	٧٥	٣١	٦٥	٨٥	١٤	٩١	٠٠	٣١	٩٠	٧٩
١١	٣٤	٦٧	٦٥	٧٦	٦٣	٤٠	٢٠	٤٢	٩٤	٠٩	١٦	٠٥	٥٩	١٥	٠٠	٢٣	٠٧
٥١	٤٠	٣٨	١٩	١٤	٠٠	١٩	٣٣	٣٢	٣٠	٤٦	٩٥	٥١	٠١	٢٤	١٤	٠٨	٩٠
٣٠	٥٦	٩٧	٠٠	٣٠	٦٥	٨٥	١٢	٠٣	٤١	١٣	٣٤	٨٢	٢١	٠٢	٦٢	٨٢	٥٣
٦٤	٥٩	٣٣	٢٣	٣١	٣٩	٨٤	٥٤	٣٣	٢٠	٢٥	٧٦	٥٠	٠٤	١٥	٢٦	١٧	٩٨
٤٤	٥٩	٣٥	٨٩	١٧	٦٥	٥٤	٦٤	٥٠	٤٥	٦٢	٣٠	٤٠	٨٣	٤٤	١١	٩١	٠٨
٥٠	٩٠	٧٤	١١	٤١	٣٣	٣٧	٩٧	٥٤	٨٥	١٩	٦٧	٨٧	٨٤	٧٧	٤٦	٢١	٣٧

وتعتمد طريقة استخدام جداول الأرقام العشوائية على حجم العينة المطلوب سحبها من المجتمع الأصلي. فعند إستخدام هذه الجداول تعطى أرقام مسلسلة مفردات المجتمع الذى نريد معاينته، ثم يختار اختياراً عشوائياً بداية تؤخذ من عندها الأرقام التى تدخل ضمن مفردات العينة مع استبعاد الأرقام المتكررة أو الأرقام التى تزيد عن حجم المجتمع. ولتحقيق ذلك يلزم أخذ أعمدة تحتوى على عدد من الخانات تساوى عدد أرقام حجم المجتمع، ولشرح ذلك نقول: لو كانت هناك بيانات عن مجتمع يتكون من ١٠٠ مفردة فبإمكاننا إختيار الأرقام بالطريقة العشوائية من العمودين الأول والثاني من الجدول رقم (١-١). فإذا كنا نريد اختيار عينة مكونة من ١٠ مفردات من ١٠٠ مفردة، فإننا سنرى أن أرقام المفردات المطلوبة، من الجدول السابق، هى ٢٠، ٧٤، ٤٩، ٢٢، ٩٣، ٤٥، ٤٤، ١٦، ٤، ٣٢، أما إذا كان حجم المجتمع كبيراً كأن يتكون من ١٠٠٠٠ مفردة، ففى هذه الحالة تستعمل الأعمدة الأربعة فى الجدول المذكور. فإذا أردنا سحب عينة مكونة من ١٠٠ مفردة من هذا المجتمع، فإن أرقام مفردات العينة المختارة من الجدول بهذه الطريقة تكون هى على التوالى ١٧٢٠ ثم تليها ٤٩٧٤. وبعدها ٧٠٩٤ وهكذا حتى يصل عدد المفردات ١٠٠، وهى عدد مفردات العينة المطلوبة. وإذا كنا بصدد معاينة ٢٠٠٠ مفردة فقط فإننا نختار من الجدول السابق أرقام مكونة من أربعة حدود لمفردات العينة المطلوبة وهى ١٠٠ مفردة.

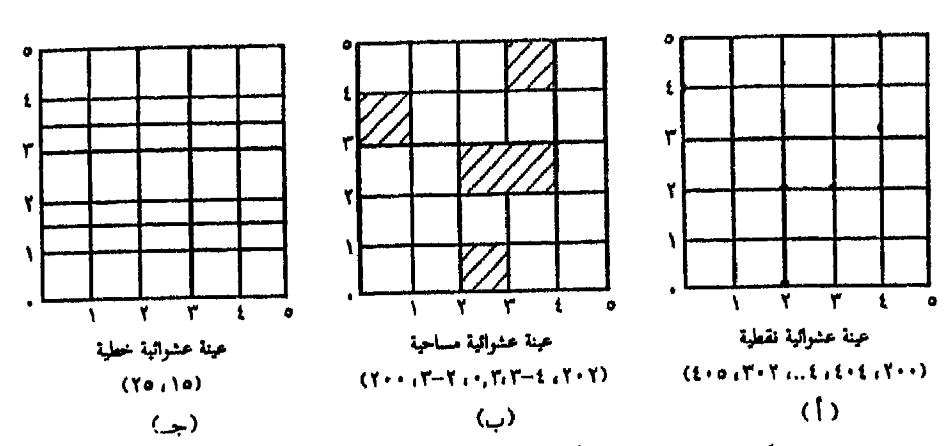
وتتم عملية الاختيار بإحدى الطريقتين: أما أن تختار الأرقام العشوائية من العمودية الأول والثانى لمفردات العينة بقبول أى رقم يقع بين ١ و ٢٠٠٠ ورفض أى رقم آخر يقع بين ٢٠٠٠و ٩٩٩٩، ونستمر فى عملية القبول والرفض حتى يصل عدد المفردات ١٠٠ مفردة. ويعاب على هذه العملية أنها تستنفذ وقتاً طويلاً وذلك بسبب ارتفاع معدل رفض الأرقام التى تزيد عن الحجم الكلى للمفردات الذى نريد سحب العينة منه. ولهذا السبب تستبدل طريقة قبول ورفض الأرقام العشوائية بطريقة أخرى يمكن بها اختيار المفردات المطلوبة وذلك على أساس قبول

جميع الأرقام التي تزيد عن رقم ٢٠٠٠ واعتبارها تكرارات للأرقام بين ١ و ٢٠٠٠، بمعنى أن الأرقام العشوائية بين ٢٠٠١و ٤٠٠٠ و٤٠٠١ و ٦٠٠٠ و ٦٠٠١ و ٨٠٠٠ و٨٠٠١ و ١٠٠٠٠ يمكن اعتبارها أرقاماً إضافية (جديدة) للأرقام من ١ إلى ٢٠٠٠. وتمتاز هذه الطريقة بأن فيها اختصار كبير للوقت في إختيار أرقام المطلوبة من جميع الأرقام المحصورة بين ١ و ١٠٠٠٠.

وفي بعض الحالات تكون البيانات المتاحة عن مجتمع الظاهرة قيد البحث في شكل مجموعات ويراد أخذ عينة عشوائية من المجموعات ككل وليس من كل مجموعة على حدة. فمثلاً إذا كانت لدينا بيانات عن أعداد السكان لعدد كبير من المراكز العمرانية التي تتكون من عدة مجموعات: قرى صغيرة، قرى كبيرة، مدن صغيرة، مدن كبيرة (عواصم المحافظات)، فإنه بإمكاننا أخذ عينة واحدة من كل هذه المجموعات باستخدام الطريقة العشوائية في الاختيار. ويتم ذلك بأن ترتب أعداد كل المجموعات في قائمة حتى نحصل على المجموع الكلي لأعداد جميع المراكز العمرانية المطلوب معاينتها. فمثلاً إذا كان عدد المراكز العمرانية للمجموعات الأربع حهو ١٣٦٠، ٦٧، ٣٢، ٧ على الترتيب، فإنها تأخذ ترقيماً من ١ إلى ١٣٦، ومن ١٣٧ إلى ٢٠٤، ومن ٢٠٥ إلى ٢٣٧، ثم من ٢٣٨ إلى ٢٤٥ على الترتيب أيضاً. وبالتالي يتكون المجتمع الاحصائي للمجموعات الأربع من ٢٤٥ مركزاً عمرانياً، والذي يمكن منه سحب أية عينة بالطريقة العشوائية السابق شرحها.

وبالإضافة إلى الاختيار العشوائي لمفردات العينة من بين مفردات المجتمع التي تحتويها قائمة بأعداد هذه المفردات، فإنه يمكن أيضاً تطبيق نفس أسلوب الاختيار على البيانات الخاصة بالتوزيعات المكانية Spatial (Areal) Distribution التي تعد حجر الزاوية في الدراسات الجغرافية. ويتطلب إجراء معاينة التوزيعات المكانية بطريقة الاختيار من الجداول العشوائية أن يرسم على خريطة المنطقة قيد البحث شبكة من المربعات، في حالة إذا لم يوجد على الخريطة شبكة خطوط الطول ودوائر العرض، ويعطى لخطوط شبكة المربعات (الطولية والعرضية) التي أنشأناها أرقاماً أو إحداثيات

من الغرب إلى الشرق ومن الجنوب إلى الشمال بحيث يكون رقم الصفر (بداية التقسيم أو الترقيم) في الركن الجنوبي الغربي (نقطة الأصل) للخريطة الذى تتزايد منه الاحداثيات شرقاً وشمالاً. ويمكن وضع الأرقام بحيث تدل على خطوط الإحداثيات نفسها أو تشير إلى المسافة بين كل خط إحداثى وآخر يليه، ويتوقف ذلك على نوع التوزيعات المكانية المطلوب معاينتها. فمثلاً عند الدراسة الجغرافية لبعض مظاهر النشاط الزراعى فى منطقة ما، كدراسة الدورات الزراعية والتى يكون الهدف منها معرفة التأثير الذى يفرضه استخدام دورة زراعية معينة على الإنتاج الزراعى (مع افتراض ثبات تأثير العوامل الأخرى)، فإننا فى هذه الحالة نختار مواقع الأماكن الزراعية (أو القرى وزمامها الزراعى) كإطار للمعاينة لأنه من المعروف أن تلك الأماكن تمثل المصدر الذى يستمد منه البيانات الخاصة لمثل هذه الدراسة. وبما أننا نريد اختيار مجموعة من القرى على خريطة كعينة للدراسة فإننا نكون بصدد تطبيق أسلوب المعاينة النقطية Point - Sampling والذى يمكن اعتباره أفضل الطرق لهذا الغرض. ولتحديد مواقع القرى المطلوبة على خريطة لاتتعدى فيها خطوط الإحداثيات عن ١٠٠ خط فى كل إتجاه على حدة، نقوم بإختيار العمودين الأول والثانى أو أى عمودين آخرين من جدول الأرقام العشوائية (جدول رقم ١ - ١)، ونمثل برقمى العمود الأول المسافة على الإحداثى الشرقى (المحور الأفقى العشوائى) من نقطة الأصل، وبرقمى العمود الثانى المسافة على الإحداثى الشمالى (المحور الرأسى العشوائى) من نقطة الأصل أيضاً. أو بعبارة أخرى تدل الأرقام العشوائية المختارة على قيم أو أرقام خطوط الإحداثيات لكل محور على حدة. وعلى أساس الإحداثى (الشرقى والشمالى) العشوائى المستخرج من الجدول يرسم عمودان أحدهما على المحور الأفقى العشوائى والآخر على المحور الرأسى العشوائى حيث تحدد نقطة تلاقيهما موقع القرية التى تدخل ضمن مفردات العينة المطلوبة (شكل رقم ١ - ٢أ).



شكل رقم (١-٢): أنواع الاختيار العشوائى للتوزيعات المكانية

أما إذا كنا نريد معاينة مظاهر استخدام الأرض Land Use فى منطقة ما، فإننا نكون بصدد تطبيق أسلوب المعاينة المساحية Area - Sampling والذى هو أفضل الأساليب لمثل هذه المعاينة. ويتطلب إجراء المعاينة المساحية على خريطة إما أن ترقم المسافات الإحداثيات (شكل رقم ١ - ٢ب) بأرقام عشوائية تستخرج من الجدول وتحدد المساحات المطلوبة للعينة داخل المربعات بطريقة الإحداثيات السابق ذكرها، وإما أن نبقى على ترقيم خطوط الإحداثيات كما هى الحال فى طريقة المعاينة النقطية (شكل رقم ١ - ٢ب) وتختار المربعات (أو المساحات) المطلوبة بطريقة الإحداثيات أيضاً.

ومن البيانات الجغرافية ما يختص بالتوزيعات المكانية الخطية (مثل خطوط أو شبكات المواصلات بأنواعها المختلفة، والأنهار وروافدها، أو أى ظاهرة جغرافية تتخذ شكل الامتداد الخطى) والتى يمكن معاينتها بالمعاينة الخطية Line - Sampling. وفى هذا الصدد فإن إجراء سحب عينة عشوائية خطية لايتطلب سوى تحديد الإحداثيات على أحد المحورين، أو كليهما، عن طريق اختيار أرقامها من عمود واحد أو أكثر من جدول الأرقام العشوائية لتمثيل مفردات العينة المطلوبة. وعلى طول خط الإحداثى الذى حددناه عشوائياً يمكن قياس المسافات التى تشتمل - أو لاتشتمل - على خصائص معينة للظاهرة قيد البحث والتى تمثل قيمها مكوناً أساسياً للبيانات التى سيجرى عليها التحليل الكمى بعد ذلك.

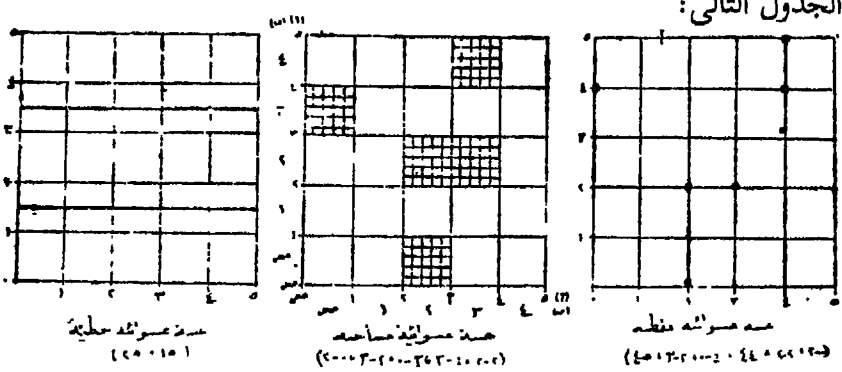
وجدير بالذكر أننا فى كل الحالات السابقة كنا نفترض أن المساحة التى تجرى عليها أحد طرق المعاينة تأخذ شكلاً مساحياً مربعاً أو مستطيلاً يمكن رسم شبكة متقنة ودقيقة من المربعات داخله، ولكن ليس هذا بالوضع الحقيقى فى الطبيعة لمعظم التوزيعات المكانية والتى قد تمثل مناطق أو تقسيمات إدارية تتخذ على الخريطة أشكالاً مساحية مختلفة تبعاً لتباين شكل حدودها. وللتغلب على هذه الصعوبة فإنه بإمكاننا استخدام مستطيل من ورق الشفاف، مرسوم عليه شبكة دقيقة من المربعات، تفوق مساحته المنطقة كلها أو يغطى معظمها على الأقل. وبذلك يمكن رفض أى إحداثى يختار من جدول الأرقام العشوائية ويقع بعيداً عن حدود المنطقة قيد البحث.

حـ- طريقة السحب بواسطة الحاسب الآلى:

تستخدم هذه الطريقة فى حالة سحب عينات كبيرة من الحجم من مجتمعات تتميز بأحجام كبيرة جداً. وتحقق هذه الطريقة درجة عالية من العشوائية وعدم التحيز حيث أن الباحث لايتدخل فى عملية الاختيار.

ويطبق هذا الأسلوب حالياً فى الأبحاث العلمية التى تجرى فى معظم دول غرب أوربا والولايات المتحدة الأمريكية، وفى سبيله للإنتشار فى الدول الأخرى التى أخذت على عاتقها، حديثاً، تطوير أجهزتها الإحصائية بإدخال الحاسبات الآلية Computers ومن بينها جمهورية مصر العربية.

وبالإضافة إلى السحب الآلى لمفردات العينة، فإنه من الممكن الآن تغذية الحاسب الآلى ببرامج يمكن بواسطتها أن يضع جداول للأرقام العشوائية التى من أمثلتها الجدول التالى:



جدول رقم (١ -٢) جدول الأرقام العشوائية بواسطة الحاسب الآلى(١)

٧ ٣ ٢ ٤ ٤ ٧ ٠ ١ ١ ٤ ٠ ٧ ٠ ٤ ٤ ٠ ١ ٢ ٤ ٤ ٥ ٥ ٩ ٩ ٠ ٣ ٣
١ ١ ٣ ٢ ٤ ٣ ٩ ٥ ٢ ٤ ٠ ٦ ٢ ١ ٣ ٢ ٠ ٩ ١ ٤ ١ ٣ ٣ ٧ ٥ ٧ ٨ ٠
٧ ٦ ٢ ٤ ٥ ٢ ٣ ٢ ١ ٣ ٤ ٣ ٧ ٦ ٨ ٩ ٨ ٨ ٥ ٨ ٤ ٤ ٦ ٢ ٣ ٨ ٣ ٤
٣ ٠ ٩ ٠ ٧ ٤ ٣ ١ ٩ ٢ ٣ ٣ ٣ ٧ ١ ٩ ٥ ٣ ١ ١ ٤ ٢ ٨ ٤ ٤ ٢ ٨ ٤
٢ ٩ ٢ ٦ ٢ ٣ ٤ ٧ ٢ ١ ٣ ٨ ٥ ٦ ١ ٤ ٢ ٠ ١ ٢ ٢ ٨ ٤ ٦ ٣ ٠ ٧ ٢
٠ ٤ ١ ٤ ١ ٩ ٧ ٧ ٨ ٦ ٤ ٣ ٨ ٠ ٨ ٩ ٤ ٤ ٤ ٧ ٢ ١ ٢ ١ ٣ ٣ ٢ ٢
٤ ٣ ٨ ٦ ١ ٨ ٤ ٥ ٣ ٢ ٤ ٠ ٣ ٠ ٤ ٦ ١ ٨ ٣ ٧ ٨ ٣ ٤ ٣ ١ ٨ ٥ ٦
٢ ٠ ٣ ٧ ٠ ٢ ٧ ١ ٣ ٩ ٨ ٨ ٥ ٨ ٧ ٩ ٧ ١ ٧ ٤ ٠ ٣ ٠ ٨ ٤ ٩ ٦ ١
٧ ١ ٤ ٩ ٦ ٥ ٥ ٥ ٩ ٢ ٩ ٥ ٦ ٨ ٢ ٢ ٧ ٢ ٤ ٦ ٣ ٤ ٦ ٦ ٧ ٧ ٨ ٧
٩ ١ ٧ ٢ ٤ ٩ ٩ ٤ ٧ ٣ ١ ٧ ٤ ٦ ٣ ٣ ٨ ٦ ٧ ٢ ٠ ٢ ٩ ١ ٤ ٥ ٣ ٩
٣ ٨ ٠ ٤ ٣ ٠ ٣ ٥ ٩ ٧ ٧ ٤ ٢ ٠ ٦ ١ ٥ ٩ ٩ ٥ ١ ٢ ١ ٤ ١ ٩ ٥ ٤
٦ ٩ ٥ ٦ ٥ ٤ ٠ ٤ ١ ٠ ٩ ٠ ٩ ٩ ٥ ١ ٦ ١ ٥ ٨ ٦ ٠ ٠ ٣ ٨ ٨ ١ ٨
٣ ٧ ٨ ٥ ٣ ٦ ٤ ٠ ٢ ٢ ٨ ٤ ٨ ٩ ٢ ٩ ٠ ٣ ٥ ٤ ٧ ٩ ٢ ١ ٥ ٣ ٧ ٠
٠ ٥ ٦ ٧ ٧ ٩ ٥ ٢ ٩ ٤ ٢ ٧ ٤ ٠ ٢ ٤ ٣ ٣ ٢ ٥ ٣ ٣ ٧ ٥ ٠ ١ ٢ ١
٣ ٧ ٩ ٦ ٥ ٤ ٣ ٤ ٩ ٦ ٠ ٥ ٨ ٣ ٠ ٧ ٥ ٧ ٦ ٨ ٩ ٠ ٧ ٨ ٣ ٣ ٩ ٢
٩ ١ ٢ ٤ ٥ ١ ٠ ٩ ٤ ٨ ٥ ٠ ٤ ٩ ٧ ٩ ٩ ٦ ٠ ٧ ٠ ٦ ٤ ٨ ٩ ٩ ٤ ٨

(١) وضع برنامج الحاسب الآلى للحصول على أرقام هذا الجدول Dr. M. McCullagh أستاذ الكارتوجرافيا والتحليل الكمى فى الجغرافيا - بقسم الجغرافيا - جامعة نوتنجهام - انجلترا.

٦ ٧ ٧ ٢ ٤ ٤ ٠ ٢ ٠ ٦ ١ ٠ ٥ ٩ ١ ٦ ٨ ٤ ٢ ٠ ٠ ٤ ٩ ٢ ٥ ٨ ٢

١ ٨ ٩ ٧ ٦ ٨ ٨ ٨ ٧ ٩ ٥ ٩ ٩ ٩ ٦ ٦ ٠ ٥ ٥ ١ ٧ ٥ ٦ ٧ ٦ ٦ ٢ ٧

٦ ٢ ٩ ٩ ٤ ٣ ٤ ٢ ٩ ٢ ٢ ٥ ٣ ٦ ٧ ٢ ٠ ٥ ٩ ٥ ٥ ٥ ٦ ١ ٩ ٥ ٣ ٠

٧ ٥ ٦ ٧ ٩ ٦ ٦ ٢ ٧ ٧ ٦ ٧ ٧ ٢ ٤ ٤ ٩ ١ ١ ٠ ٦ ٨ ٧ ٦ ٧ ٨ ٧ ٨

٥ ٥ ٦ ١ ٣ ٩ ٥ ٣ ٠ ٢ ١ ٨ ٩ ٧ ٦ ٨ ٨ ٨ ٧ ٩ ٥ ٩ ٩ ٩ ٦ ٠ ٥ ٥

٦ ٨ ٧ ٦ ١ ٧ ٨ ٧ ٨ ٧ ٦ ٢ ٩ ٩ ٤ ٣ ٤ ٢ ٩ ٢ ٢ ٥ ٣ ٦ ٢ ٠ ٥ ٩

٦ ٠ ٨ ٠ ٥ ٨ ٠ ٥ ٥ ١ ٧ ٥ ٦ ٧ ٩ ٦ ٦ ٢ ٧ ٧ ٦ ٧ ٧ ٢ ٤ ٩ ١ ١

٤ ٥ ٧ ٣ ٤ ٩ ١ ١ ٤ ٨ ٦ ٨ ٥ ٨ ٩ ٣ ٧ ٧ ٧ ٠ ٢ ٠ ٦ ١ ٨ ٢ ٥ ٤

٨ ٢ ٧ ٢ ١ ٤ ٩ ٦ ٤ ٨ ٤ ٩ ٨ ١ ٤ ٥ ٢ ٧ ٦ ٥ ٦ ٦ ٧ ٧ ٠ ١ ٢ ٦

٩ ٣ ٧ ٠ ٨ ١ ٩ ٤ ٥ ٩ ٤ ٥ ٥ ٧ ٥ ٩ ٨ ٠ ٧ ٥ ٧ ١ ٤ ٦ ٨ ٠ ٦ ٠

٢ ١ ٣ ٨ ٩ ٥ ١ ٢ ٦ ٧ ٧ ٤ ٧ ٩ ٩ ١ ٣ ٧ ٢ ٥ ٠ ٨ ٣ ٤ ٩ ٦ ٠ ٣

٨ ٣ ٠ ٣ ٦ ٢ ٦ ٦ ٢ ٦ ٦ ٢ ٠ ٣ ٠ ٧ ١ ٢ ٢ ٢ ٣ ٩ ٢ ٩ ٤ ٤ ٦ ٢

٦ ٥ ٦ ٩ ٥ ١ ٩ ٩ ٢ ٢ ٤ ٦ ٣ ٦ ٨ ٣ ٢ ٣ ١ ٧ ٧ ١ ٣ ٣ ٨ ٧ ٢ ٩

٩ ٩ ٠ ١ ٧ ٥ ٣ ٩ ٢ ٣ ٠ ٢ ١ ٣ ٤ ٦ ٣ ٢ ٤ ٩ ٤ ٥ ٦ ١ ٥ ٠ ٩ ٦

٢ ٠ ٢ ٠ ٢ ٧ ٥ ١ ٩ ٧ ٥ ٣ ٦ ٨ ٥ ٢ ٧ ٢ ١ ٠ ٨ ٤ ٢ ٢ ٣ ٨ ٣ ٥

٨ ٣ ٧ ٨ ١ ١ ٩ ٧ ٦ ٩ ٦ ٣ ١ ٤ ٢ ٣ ٤ ٣ ٩ ٠ ٠ ٤ ٠ ٧ ٤ ٣ ١ ٤

٥ ٠ ٦ ٦ ٥ ٥ ٦ ٣ ٨ ٠ ٨ ٤ ١ ٦ ٥ ٩ ٧ ١ ٥ ٢ ٣ ٥ ٢ ٩ ٩ ٧ ١ ٨

٠ ٦ ٢ ٧ ٥ ٢ ١ ٠ ٣ ٣ ٨ ٣ ٢ ٦ ٤ ٢ ٠ ٣ ٢ ٣ ٧ ٤ ٥ ٠ ٠ ٦ ٨ ١

٦ ٩ ٠ ٨ ١ ٣ ٦ ٥ ٥ ٩ ٠ ١ ٢ ٣ ٦ ٧ ٩ ٠ ٣ ٥ ١ ٦ ٩ ٧ ٣ ١ ٥ ٨

نخلص من كل مما سبق أنه عل يالرغم من أن الباحث يجب أن يكون حذراً وغير متحيزاً عند اختياره لمفردات العينة بإحدى طرق الاختيار السابقة، إلا أن هناك أنواع كثيرة من الأخطاء التى تصاحب أسلوب سحب العينات يكون مصدرها الرئيسى إما تحيز الباحث (خطأ التحيز Bias Error) أو تحيز المبحوث، أو عدم التزام الباحث بأسلوب العمل الإحصائى أى عدم اتباع القواعد السليمة فى جمع المعلومات أو سوء التقدير والإهمال فى العمل. وإلى جانب ذلك هناك أيضاً خطأ الصدفة أو الخطأ العشوائى Random Error الذى يعد من أهم أخطاء أسلوب المعاينة فى الدراسة. وكما سبق أن ذكرنا أن خطأ الصدفة من الأخطاء التى تخرج عن نطاق القصد أو التعمد حيث أن قيمة الخطأ تتفاوت من عينة لأخرى، وأن مصدره الأساسى هو الاعتماد على بيانات العينة فقط فى استخلاص النتائج الخاصة بالمجتمع الذى تمثله هذه العينة (راجع الفصل الأول).

مثال تطبيقى لكيفية سحب عينة:

لنفرض أنه لدراسة تباين الإنتاج الزراعى فى إحدى المحافظات عن طريق اختبار الفرق بين متوسطات إنتاج الأرض الزراعية فى قرى تلك المحافظة والتى يبلغ عددها ٩٥ قرية، ولما كان من الصعب دراسة القرى كلها لكثرة عددها أو لصعوبة الوصول إلى بعض منها، فقد تقرر أخذ عينة من ١٠ قرى. والمطلوب تحديد هذه القرى العشرة من بين القرى كلها. اولتحقيق ذلك تستخدم طريقة الجداول العشوائية لاستخراج أرقام القرى المطلوبة باتباع الخطوات التالية.

١- إعطاء كل قرية رقماً خاصاً بها من ١ إلى ٩٥.

٢- بالرجوع إلى جداول الأعداد العشوائية (جدول رقم ١-١) يمكن اتخاذ أى عمود أو صف واختيار عشرة أرقام منه.

٣- إذا أخذنا الصف الأول من الجدول فإن الأرقام العشوائية للقرى المختارة تكون هى:

٢٠، ١٧، ٤٢، ٢٨، ٢٣، ٥٩، ٦٦، ٣٨، ٦١، ٢ (لاحظ أننا لم نأخذ

الرقم السادس فى الصف وهو رقم ١٧ لأنه مكرر واستبدلناه بالرقم ٢ من نفس الصف).

٤- تكون الأرقام العشرة السابقة هى الأرقام العشوائية الممثلة لمجتمع القرى (٩٥).

أما إذا كان عدد القرى هو ٩٥٠ قرية وأريد اختيار عينة بحجم ما فإن أرقام مفردات العينة ستحتوى فى هذه الحالة على أرقام مكونة من ثلاثة حدود. فإذا كانت العينة مكونة من ٣٠ قرية مثلاً فإننا نأخذ العمودين الأول والثانى من جدول الأرقام العشوائية ونختار منها الأرقام المكونة من ثلاثة حدود، والتى تمثل مفردات العينة، فتكون هى:

٧٢٠، ٠٩٤، ٥٢٢، ٤٤٥، ١٤٤، ٣١٦، ٠٠٤، ٠٣٢، ٤٠٣، ٠٦١، ٣٨٩، ٢٠١، ٦٢٧، ٥٤٩، ٤٤٩، ٦٢٠، ٧٤٨، ٢٠٨، ٧٩٥، ٩٣٧، ٩٠٥، ٥٥٥، ٨٦٧، ٦٨٥، ٠٤٠، ٥٩٤، ٣١١، ٠٦٤، ٤٥٠، ٨٨٦.

ويلاحظ أننا رفضنا الأرقام ٩٧٤، ٩٩٣، ٩٦٢ نظراً لأنها أكبر من حجم المجتمع الأصلى (٩٥٠ قرية) واخترنا بدلاً منها الأرقام العشوائية الثلاثة الأخيرة: ٠٦٤، ٤٥٠، ٨٨٦.

(٣) تحديد نوع العينة:

يجمع كثير من الحصائيين والباحثين على أن تحديد نوع العينة المختارة، التى يجب أن تتوفر فيها صفة إعطاء نتائج ذات دقة معينة بأقل تكاليف ممكنة أو بأعلى دقة بتكاليف محدودة، يتوقف على طبيعة الدراسة، ونوعية وتركيب المجتمع الذى ستسحب منه العينة، والوسيلة أو الأداة المستخدمة فى جمع البيانات، ووجهة نظر الباحث نفسه.

ويمكن تصنيف العينات على أساس عامل العشوائية فى الاختيار إلى قسمين رئيسيين: القسم الأول يشمل العينات العشوائية التى يعتمد الباحث فى تصميمها على نظرية الاحتمالات فى إعطاء الفرص المتكافئة لمفردات المجتمع لأن تظهر فى العينة، أما القسم الثانى فيتضمن العينات العمدية (غير العشوائية) والتى يكون فيها

تحيز البحث واضحاً فى اختيار مفردات العينة وذلك عن طريق إعطاء فرص غير متكافئة للمفردات نتيجة تعمده اختيار بعض المفردات دون غيرها من مفردات المجتمع الذى يريد معاينته. ولكل من القسمين أنواع متعددة ومتنوعة من العينات سندرسها بالتفصيل كما يلى:

أولاً: العينات العشوائية Random Samples

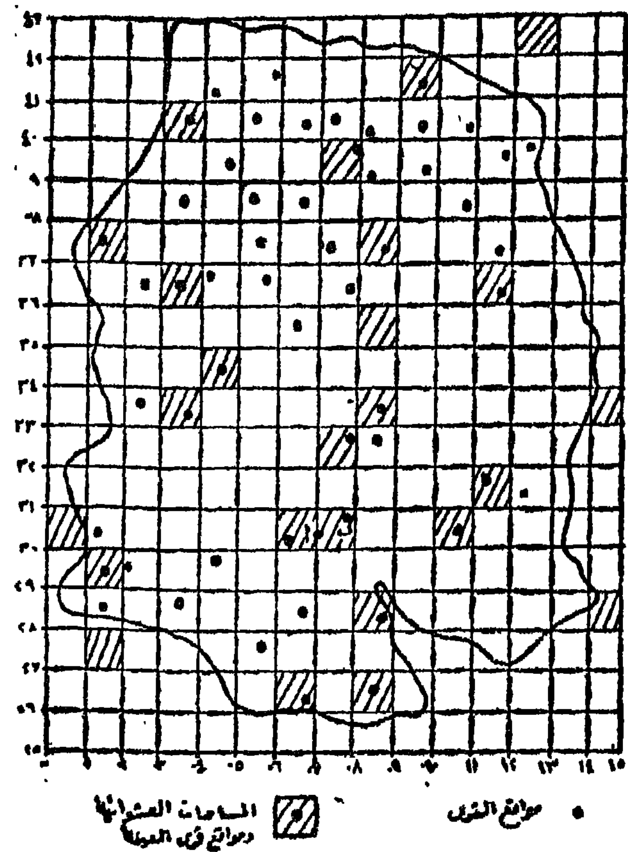
يستخدم أحياناً مصطلح العينات الاحتمالية للدلالة على العينات العشوائية وذلك لأنها تعتمد، كما سبق القول، على نظرية الاحتمالات فى اختيار مفردات العينة من بين مفردات المجتمع عن طريق سحب تلك المفردات بالتتابع لكل منها احتمال معروف فى الاختيار فى السحبات المختلفة. وبعبارة أخرى تجرى المعاينة العشوائية (الاحتمالية) حسب خطة إحصائية لايسمح فيها للباحث أو حتى لمفردات العينة التدخل فى اختيار العناصر الخاصة بالبيانات المراد جمعها أو أن يستعاض عن بعض المفردات التى يجب اختيارها وضمها للعينة بمفردات أخرى أسهل فى حالة صعوبة الوصول إلى أو الحصول على بيانات المفردات الأولى. وسوف نعرض فيما يلى لأهم أنواع العينات العشوائية وهى: العينة العشوائية البسيطة، العينة العشوائية المنتظمة، العينة العشوائية الطبقية، والعينة العشوائية المتعددة المراحل.

أ- العينة العشوائية البسيطة Simple Random Sample

يلائم هذا النوع من العينات الدراسات التى تهدف إلى تحديد خصائص المجتمع الذى تمثل مفرداته مجموعة متجانسة، أى ذات نوعية واحدة، مثل مجموعة من الطلاب عند مستوى عمرى متقارب وأوزان متساوية تقريباً، أو مثل إنتاج أحد مصانع المعلبات لوحدات إنتاجية (عبوات) ذات أوزان واحدة. ويتم إختيار مفردات العينة على أساس تساوى فرص الاختيار المستقل أمام كل مفردات المجتمع. مثال ذلك إذا أردنا إختيار عينة من ٢٠ قرية فى منطقة تحتوى على ٢٠٠ قرية لتكون فى قرى العينة صفات تمثل جميع القرى فى المنطقة، فإنه يجب اختيارها بالطريقة العشوائية التى عرفناها. كما يمكن بطريقة أخرى تحديد عدد

قرى العينة بتعيين مواقعها على الخريطة باستخدام الإحداثيات العشوائية (شكل رقم ١ - ٣). وفيما يلى أهم الخطوات التى تتبع لإجراء هذا الاختيار.

١- نبدأ بتحديد إطار المعاينة الذى يؤكد المسافات العشوائية بين القرى عن طريق استخدام مستطيل شبكة المربعات الذى يوضع فوق الخريطة، وتكون أماكن القرى المطلوبة ما يقع منها داخل المربعات التى تختار بطريقة الإحداثيات العشوائية. وفى هذه الحالة تقوم بترقيم الإحداثيات الرئيسية لشبكة المربعات، باختيار أرقام من جدول الأرقام العشوائية، ابتداء من الركن الجنوبى الغربى (نقطة الأصل) للخريطة. وبما أنه من المتوقع إختيار إحداثيات لمربعات خالية من مواقع القروى داخل المنطقة، أو مربعات يوجد بها قرى ولكنها تقع خارج



شكل رقم (١ -٣): تحديد مواقع ٢٠ قرية باستخدام الاحداثيات العشوائية

حدود المنطقة قيد البحث، فإنه من المستحسن أن نأخذ عينة من ٢٥ مربعاً ثم نختار منها ٢٠ مربعاً تحتوى على مواقع القرى المطلوبة.

٢- بالرجوع إلى جدول الأرقام العشوائية (جدول رقم ١-١) نختار أى زوجين من الأرقام فى الجدول ليمثل كل زوج منها أحد الإحداثيين الشرقى أو الشمالى، وبذلك تكون الإحداثيات العشوائية لعدد ٢٥ قرية هى:

٣٦ – ١١	٣٠ – ١٠	٢٦ – ٠٨	٣٠ – ٠٦	٣٣ – ٠٧
٢٧ – ٠١	٢٨ – ١٤	٢٩ – ٠١	٣٣ – ١٤	٣٦ – ٠٣
٣٩ – ٠٧	٣٧ – ٠٨	٤٠ – ٠٣	٢٣ – ٠٣	٣٧ – ٠١
٤١ – ٠٩	٤٢ – ١١	٣٤ – ٠٤	٣٥ – ٠٨	٣٦ – ٠٦
٣١ – ١١	٢٨ – ٠٨	٣٣ – ٠٨	٣٠ – ٠٧	٣٠ – ٠٠

٣- يؤخذ الزوج الأول من الأرقام من كل عمود ليمثل الإحداثيات الشمالية لشبكة معاينة المربعات والتى يبدأ القياس على محورها من ٢٥ حتى ٤٣، بينما يمثل الزوج الآخر من الأرقام الإحداثيات الشرقية التى يبدأ القياس على محورها من صفر حتى ١٥.

٤- بطريقة الإحداثيات السابقة يمكن تحديد المربعات التى تقع بداخلها القرى المطلوبة. وتظلل هذه المربعات حتى يمكن تمييزها عن بقية المربعات التى لم يقع عليها الاختيار. ويلاحظ من الشكل وجود خمسة مربعات مظللة بدون أى موقع لقرية بداخلها. كما أنه فى حالة إذا ما وجد بداخل المربع المختار أكثر من موقع للقرى فيفضل اختيار موقع القرية القريب من الركن الجنوبى الغربى للمربع، بينما إذا كان هناك موقعاً لقرية فوق خط إحداثى يفصل بين مربعين، فيفضل ضم هذا الموقع إما إلى المربع شمال هذا الإحداثى أو إلى المربع فى شرقه. ويوضح الإحداثى (٣٠ – ٠٧) هذه الحالة والتى ضم فيها الموقع (أ) إلى المربع الذى يقع إلى الشرق من الإحداثى لأنه يقترب من الركن الجنوبى الغربى لهذا المربع، وأهمل الموقع (ب) على الرغم من وقوعه داخل المربع المختار بالإحداثى (٣٠ – ٠٧). وتجدر الإشارة إلى أن أساس مثل هذه القواعد يضعها الباحث بنفسه، كما يجب

أن نؤكد أنه ليس هناك أية خطورة من عنصر التدخل الشخصي للباحث في ذلك طالما أن القواعد الموضوعة تتبع بطريقة ثابتة وبدون تغير (أو تبديل) على كل مفردات العينة.

ب- العينة العشوائية المنتظمة Systematic Random Sample

يستخدم هذا النوع من العينات عند دراسة المجتمعات التي تكون مفرداتها متخذة شكل انتظام متسق (أي تتصف بعدم التغير أو التباين الشديد) ومتدرج من حيث التنوع. وفي هذه العينة تأكيد على تسلسل المفردات وفقاً للتنوع في الخصائص المميزة للمجتمع الأصلي. ويتبع في اختيار مفردات العينة الأسلوب العشوائي، كما في العينة العشوائية البسيطة، غير أن الاختيار يتم بطريقة منتظمة. فنقوم أولا بترتيب مفردات المجتمع عشوائياً، أي بحيث تكون الفرص المتكافئة أمام جميع المفردات للظهور في العينة هي أساس الاختيار، وعندئذ تنتهي العشوائية ويبدأ اختيار مفردات العينة وفق نظام أو قاعدة معينة حتى نحصل على العينة المطلوبة.

وتمتاز هذه العينة بانتظام وثبات الفترات أو التباعد بين وحدات أو مفردات العينة. وفيها نبدأ بتحديد مقدار تمثيل كل مفردة من مفردات العينة لعدد معين من مفردات المجتمع، ثم نقوم بإختيار المفردة الأولى (أو أحد أعداد مقدار أو نسبة التمثيل) عشوائياً ونضيف إليها مقدار التمثيل بطريقة منتظمة حتى نحصل على بقية مفردات العينة بشكل منتظم وبتباعد متساوى فيما بينها. ولتحديد مقدار التمثيل نقسم عدد مفردات المجتمع الأصلي على حجم العينة المطلوب. فمثلاً إذا كان حجم المجتمع الأصلي ٦٠٠٠ مفردة، وأردنا إختيار عينة من ٢٠٠ مفردة، فهذا يعني أننا نريد اختيار مفردة واحدة لكل ٣٠ مفردة (أي $\frac{٦٠٠٠}{٢٠٠}$ = ٣٠).

وإحدى الطرق المتبعة هي: أن نختار عدداً عشوائياً بين ١ ، ٣٠. ولو فرضنا أن هذا الرقم هو ٢٢ ، فإنه يمكن تحديد مفردات العينة مباشرة بإضافة مقدار التمثيل (٣٠) بطريقة منظمة إلى الرقم (٢٢) وذلك على النحو التالي:

٢٢، ٥٢، ٨٢، ١١٢، ١٤٢، ١٧٢، ٢٠٢، ٢٣٢، ٢٦٢، ٢٩٢ ... وهكذا حتى نصل إلى المفردة الأخيرة ويكون رقمها ٥٩٩٢. وهناك طريقة لإيجاد ترتيب أى مفردة من مفردات العينة وذلك باستخدام المعادلة الآتية:

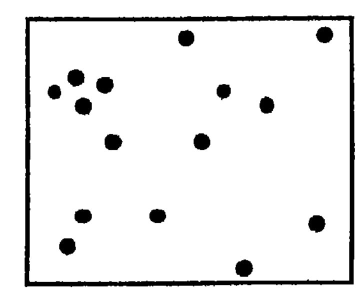
ترتيب المفردة = الرقم العشوائى المختار + (حجم العينة – ١) × مقدار التمثيل

أى أن:

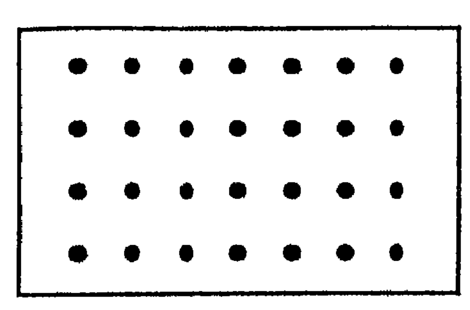
فإذا أردنا مثلاً معرفة ترتيب المفردة رقم ١٠٠ من ٢٠٠ مفردة فى العينة السابقة فإن.

١٠٠ = (٢٢ + (٢٠٠ – ١) × ٣٠) = ٢٩٩٢

ويمكن تطبيق العينة المنتظمة فى مجال الدراسة الجغرافية، إذ يختار موقع المفردة الأولى عشوائياً ويوقع على الخريطة بالنسبة لكل من إحداثيها الشرقى والشمالى من نقطة الأصل للخريطة، ثم تختار مواقع بقية المفردة الأخيرة بشكل منتظم يمثل تباعد (مسافة) بالنسبة للمحورين العشوائيين الأفقى والرأسى كما فى الشكل رقم (١–٤)



١– عينة عشوائية بسيطة





٢– عينة عشوائية منتظمة

شكل رقم (١ – ٤): التوزيع المكانى لمفردات العينة العشوائية

وتتميز العينة المنتظمة بأنها أسهل وأسرع فى التطبيق من العينة العشوائية البسيطة، إذا أنها لا تحتاج إلى اختيار كل المفردات بالطريقة العشوائية والذى قد ينتج عنه تكرار سحب بعض المفردات. كما أنها تمثل توزيعاً متسقاً (منتظماً) Uniform للمجتمع الذى تسحب منه العينة بعكس العينة العشوائية البسيطة التى

ينتج عنها فى معظم الأحيان توزيعات مكانية تتخذ أشكالاً عنقودية Clusters or Bunching تتباعد فيها المفردات بعضها عن بعض تباعداً لايمكن تقليله إلا بزيادة حجم العينة. فمثلاً إذا أراد أحد الباحثين تقدير نسبة الأراضى الزراعية من خريطة منطقة ما (شكل رقم ١- ٤أ)، وكانت أكبر المساحات للأراضى الزراعية تتركز بالقرب من المكان (أ) فإن قيمة النسبة فى هذه الحالة ستكون أعلى من القيمة الحقيقية لها.

ويعاب على العينات المنتظمة إذا أجريت على التوزيعات المكانية أنها تؤكد صفة الانتظام والاتساق فى الظاهرة أو الظاهرات الجغرافية موضع المعاينة والتى تكون غالباً، إن لم يكن دائماً، فى حالة تغير وتطور مستمرين. كما قد تتصف المعاينة المنتظمة بأنها لاتعطى عينة غير متحيزة أو ممثلة للمجتمع الذى سحبت منه بسبب خطأ التحيز وعدم اتباع أسلوب تكافؤ الفرص فى اختيار مفردات العينة إلا إذا كانت مفردات المجتمع الأصلى موزعة توزيعاً عشوائياً - وكثير من المجتمعات الاحصائية للظاهرات الجغرافية لاتتصف بصفة العشوائية فى توزيعاتها. وعلى العموم فإن خطأ التحيز إن وجد فى بيانات العينة المنتظمة فإن قيمته تكون صغيرة جداً بحيث يمكن إهمالها عند تطبيق أساليب التحليل الإحصائى الكمى على هذه البيانات.

جـ- العينة العشوائية الطبقية Stratified Random Sample

يستخدم هذا النوع من العينات العشوائية فى دراسة المجتمعات التى تتميز بتباين نوعيات مفرداتها والتى يمكن تقسيمها إلى مجموعات (أو طبقات Strata) لكل مجموعة أو طبقة منها خصائص ومميزات معينة، ولكنها تتباين فيما بينها تبايناً عظيماً فى هذه الخصائص والمميزات. كما يتسم هذا النوع من العينات بأنه أكثر دقة فى الاختيار العشوائى من العينات العشوائية البسيطة حيث أن المفردات الكلية للعينة الطبقية تكون أكثر تمثيلاً لجميع مجموعات أو طبقات المجتمع الأصلى، مما يؤدى إلى تقليل فى الأخطاء عند تعميم نتائج العينة على كل المجتمع. وهكذا تقوم العينة الطبقية على أساس تقسيم المجتمع الأصلى إلى طبقات ثم نأخذ عينة

عشوائية من كل طبقة بشكل يتناسب مع مفردات أو حجم تلك الطبقة. ولهذا فإن معاينة أية طبقة تتطلب عدة إجراءات تختلف عن تلك التى تتطلبها الطبقة (أو الطبقات) الأخرى. فمثلاً فى التعدادات بالعينة التى تجرى لحصر أعداد السكان فى منازل أحد الشوارع بمدينة ما تختلف عن مثيلتها لحصر عدد العاملين فى مؤسسة أو شركة ما.

وبصفة عامة يمكن تلخيص طريقة اختيار العينة الطبقية فى الخطوات الآتية:

١- تقسيم المجتمع إلى مجموعات مميزة أو فئات فرعية (مجتمعات صغيرة) متجانسة تعرف بالطبقات Strata.

٢- تحديد نسبة مفردات كل مجموعة أو طبقة Stratum بالنسبة للعدد الكلى لمفردات المجتمع الأصلى (أى $\frac{\text{حجم الطبقة}}{\text{حجم المجتمع الأصلى}}$).

٣- تحديد عدد مفردات العينة المطلوبة من كل طبقة أو ما يعرف بالعينة الفرعية Subsample التى تتحدد عن طريق نسبة حجم كل طبقة فى المجتمع الأصلى والحجم الكلى للعينة.

٤- استخدام الأسلوب العشوائى لاختيار المفردات من كل طبقة.

مثال: أراد باحث سحب عينة حجمها ٥٠٠ عاملاً من مجتمع عمالى لأحد الصناعات يبلغ حجمه ٥٠٠٠ عاملاً وذلك حسب الحالة التعليمية، فقام بتقسيم عمال المجتمع إلى فئة من العمال الأميين وعددهم ١٠٠٠ عاملاً، وفئة من العمال الحاصلين على شهادات أقل من المتوسطة ١٥٠٠ عاملاً، وفئة من العمال الحاصلين على شهادات متوسطة وفنية وعددهم ٢٥٠٠ عاملاً. فكم من العمال من كل فئة يمكن اختيارهم حتى يحصل على الحجم الكلى للعينة، ويتم ذلك بتحديد حجم العينة الفرعية عدد المفردات المطلوبة الممثلة لمجتمع الطبقة وهو يساوى:

$$\text{حجم العينة} \times \frac{\text{حجم الطبقة}}{\text{حجم المجتمع الأصلى}}$$

بالتالى فإن:

عدد المفردات من الفئة (عمال أميين) = ٥٠٠ × $\frac{١٠٠٠}{٥٠٠٠}$ = ١٠٠ رطلاً

عدد المفردات من الفئة (عمال الشهادات أقل من المتوسطة) = ٥٠٠ × $\frac{١٥٠٠}{٥٠٠٠}$ = ١٥٠ عاملاً

عدد المفردات من الفئة (عمال الشهادات المتوسطة والفنية) = ٥٠٠ × $\frac{٢٥٠٠}{٥٠٠٠}$ = ٢٥٠ عاملاً

ويكون عدد العمال المطلوب اختيارهم من كل فئة من الفئات الثلاث هو على الترتيب ١٠٠، ١٥٠، ٢٥٠ ومجموعهم لابد وأن يساوى الحجم الكلى للعينة المطلوبة ٥٠٠ عاملاً من المجتمع الأصلى لعمال تلك الصناعة.

وتجدر الإشارة هنا إلى أن الطبقية يمكن أن تكون طبقة طولية أو عرضية (فى حالة مجتمع المساحات)، أو طبقية أفقية أو رأسية (فى حالة نوع المناطق السكنية أو مستويات الدخل أو فئات العمر). وفى مجال الدراسات الجغرافية تستخدم المعاينة الطبقية بصفة خاصة عند دراسته أو مقارنة بعض المظاهر الجغرافية لمنطقتين أو أكثر، أو عند مقارنة بيانات متغيرين (ظاهرتين) وذلك عندما يكون تباينهما غير موزع بنسب متساوية بين أرجاء المنطقة قيد البحث. ففى الحالة الأولى مثلاً إذا أراد الباحث دراسة عينة من الأحواض الزراعية بأحد المراكز الهامشية للدلتا وذلك ضمن دراسة الخصائص الطبيعية والاقتصادية للتربة، فإن ذلك يستلزم منه تقسيم أحواض المركز حسب نوع التربة إلى أنواع أو طبقات مميزة من التربة (تربة رملية، تربة رملية طينية، تربة طينية)، واختيار عينات فرعية من الأحواض الزراعية متساوية العدد من كل من نوع من أنواع التربة، بينما فى الحالة الثانية إذا أريد فحص العلاقة بين ظاهرة (أو متغير) نمو النبات الطبيعى ومستوى الارتفاع عن سطح البحر لإحدى المظاهر التضاريسية الوعرة، فإن ذلك يستلزم من الباحث أن يختار عينة مناسبة (٢٠ مكاناً مثلاً) وممثلة لكل مستوى من مستويات (نطاقات) الارتفاع

(وهى عبارة عن المساحات بين خطوط الكنتور) على حدة. على أن تجمع البيانات الخاصة بكل متغير من المتغيرين: نمو النبات ومستوى الارتفاع من أماكن العينة. ويتم ذلك عن طريق تقسيم منطقة الدراسة إلى نطاقات (مستويات) ارتفاع تنحصر بين خطوط الكنتور: صفر – ١٠٠ متر، ١٠٠ – ٢٠٠ متر، ٢٠٠ – ٣٠٠ متر وهكذا، ثم يختار بعد ذلك ٢٠ مكانا داخل نطاق أو مستوى الارتفاع الأول اختيارا عشوائياً مع رهمال الأماكن التى تزيد عن حجم العينة الفرعية داخل هذا النطاق. ويستمر إجراء نفس العمل لكل نطاق حتى تكتمل بقية النطقات ويصبح بداخل كل منها عدد المفردات ٢٠٠ مكانا) المطلوب لمعاينة المنطقة. وتعرف العينة فى هذه الحالة بالعينة الطبقية المنتظمة Stratified Systematic حيث أنه قد تم فيها اختيار عدد مفردات متساوى فى كل الطبقات.

أمثلة تطبيقية:

مثال (١) لدراسة أحد مظاهر النشاط الصناعى مثل صناعة الأحذية فى منطقة تتميز بتباين توزيع منشآت أو ورش هذه الصناعة فى مختلف المراكز العمرانية فى هذه المنطقة ، فأننا نقوم بمعاينة توزيع هذه الصناعة بطريقة المعاينة الطبقية عن طريق تقسيم المراكز العمرانية فى المنطقة حسب حجم السكان بها، إلى أربع مجموعات: قرى صغيرة، قرى كبيرة، مدن صغيرة، مدن كبيرة، ويختار عينة عشوائية من المراكز العمرانية لكل مجموعة بنسبة معاينة $\frac{١}{١٠}$ (أى $\frac{\text{حجم العينة}}{\text{حجم المجتمع}} = \frac{ن}{ن.}$ = $\frac{ن}{ن.}$) والتى تتناسب وحجم كل مجموعة من المجموعات الأربع، كما يتناسب المجموع الكلى للعينات مع الحجم الكلى لمجتمع المراكز العمرانية. وتوضح ذلك البيانات التالية:

جدول رقم (١ -٣)

طريقة حساب عينة طبقية من المراكز العمرانية

المجمـــوعة	أ- عدد الوحدات (المراكز العمرانية) فى العينة	ب- العدد الكلى للوحدات فى كل مجموعة	حـ- عدد ورش صناعة الأحذية فى كل وحدات العينة	د- متوسط ورش العينة لكل وحدة [جـ / أ]	هـ- تقدير المجموع الكلى لعدد الورش فى كل مجموعـة (ب × د)
قرى صغيرة	١٢	١٢٠	٢٤	٢	٢٤٠
قرى كبيرة	١٤	١٤٠	٣٥	٢,٥	٣٥٠
مدن صغيرة	٥	٥٠	٦٠	١٢	٦٠٠
مدن كبيرة	٢	٢٠	٨٠	٤٠	٨٠٠
المجموع	٣٣	٣٣٠	١٩٩	٦,٠٣	١٩٩٠

من البيانات السابقة الخاصة بعينة طبقية مكونة من ٣٣ مركزاً عمرانياً من العدد الكلى المقدر بنحو ٣٣٠ مركزاً عمرانياً يمكن تقدير متوسط عدد ورش صناعة الأحذية فى كل مجموعة من المجموعات الأربع، وتقدير متوسط عدد الورش فى كل مركز عمرانى على حدة، بالإضافة إلى تقدير العدد الكلى لهذه الورش التى توجد فى المراكز العمرانية لكل مجموعة على إختلاف أحجامها وفى كل المنطقة موضع الدراسة. وبالتالى فإنه عن طريق استخدام مثل هذه العينات التى تتكون من مجموعة صغيرة نسبياً من المراكز العمرانية يمكن إعطاء صورة تفصيلية عن توزيع الظاهرة قيد البحث، وتعميم نتائج التحليل الاحصائى الكمى لهذا التوزيع على جميع المراكز العمرانية فى كل أرجاء المنطقة.

مثال (٢): لدراسة تأثير طول فترة الإقامة فى المنطقة التجارية لمدينة ما على مفهوم المركز التجارى للمدينة لدى السكان وتصورهم الذهنى لهذا الجزء من

المدينة، فإن خطة الباحث فى ذلك تنحصر فى إجراء سحب عينة تكون ممثلة لجميع السكان بقدر المستطاع لقياس هذا التأثير. ويحتاج الباحث قبل إجراء عملية المعاينة أن تتوفر لديه بعض المعطيات الاحصائية عن خصائص سكان تلك المنطقة التجاريتى يمكن الحصول عليها من التعدادات السكانية أو غيرها من المصادر. فمثلاً قد تتجمع لدى الباحث بيانات تفيد أن سكان المنطقة التجارية فى المدينة موضع الدراسة والذين تنحصر أعمارهم بين ٣٠ و ٥٠ سنة ينقسمون من حيث مدة الإقامة فى المنطقة إلى ثلاث مجموعات: المجموعة الأولى تشمل السكان الذين أقاموا طوال حياتهم فى المنطقة ونسبتهم ٦٠ ٪ من جملة السكان، والمجموعة الثانية تتكون من سكان أقاموا فى المنطقة لمدة تتراوح من ٥ إلى ١٠ سنوات ونسبتهم ٢٠ ٪ من جملة السكان، والمجموعة الثالثة تتضمن أقاموا فى المنطقة لمدة ٥ سنوات أو أقل ونسبتهم ٢٠ ٪ من المجموع الكلى للسكان. وفى هذه الحالة فإنه لجمع المعلومات اللازمة لمثل هذه الدراسة، يقوم الباحث بإجراء مقابلة لعينة من السكان من كل مجموعة، كأن يختار مثلاً ١٢٠ ساكناً فى نفس فئة العمر (٣٠ - ٥٠ سنة) من الذين أقاموا فى المنطقة طوال حياتهم، ٤٠ ساكناً فى نفس فئة العمر أيضاً من بين الذين عاشوا فى المنطقة لمدة تتراوح بين ٥ سنوات و ١٠ سنوات، ٤٠ ساكناً من الذين أقاموا لمدة ٥ سنوات أو أقل فى المنطقة. وبإجراء ذلك فإن الباحث يكون قد قسم مفردات العينة على أساس مدة الإقامة إلى ثلاث طبقات يتناسب حجم مفردات كل منها مع الحجم الأصلى لنفس الطبقة من المجتمع. أما إذا أراد الباحث عمل مقارنة بين الطبقات الثلاث فإنه يقوم بأخذ ثلاث عينات فرعية Sub-Smples متساوية الحجم حتى نكون المقارنة سليمة. ويعاب على الطريقة الأخيرة عند تطبيق المعاينة الطبقية أن المجموع الكلى للمفردات المختارة قد لايكون ممثلاً Representative لجميع مفردات الظاهرة قيد البحث.

مثال (٣) : فى دراسة لاستخدام الأرض وعلاقته بمظاهر السطح فى منطقة ما، أخذت عينة عشوائية من ١٠٠ موقع من تلك المنطقة، ثم صنفت هذه المواقع بحسب فئات الاستخدام المختلفة (أراضى زراعية، حشائش، غابات، أراضى جرداء)

وعلى أساس مستويات (طبقات) الارتفاع، المتباينة، فكانت بيانات هذه العينة كالآتي:

فئات الاستخدام / مستويات الارتفاع (متر)	أراضي زراعية	حشائش	غابات	أراضي بور	حجم العينة
أقل من ١٥٠	٣	٤	١	–	٨
١٥٠ – ٣٠٠	٥	٢١	٥	٣٠	٤١
أكثر من ٣٠٠	–	٦	–	٤٥	٥١
المجموع	٨	٣١	٦	٥٥	١٠٠

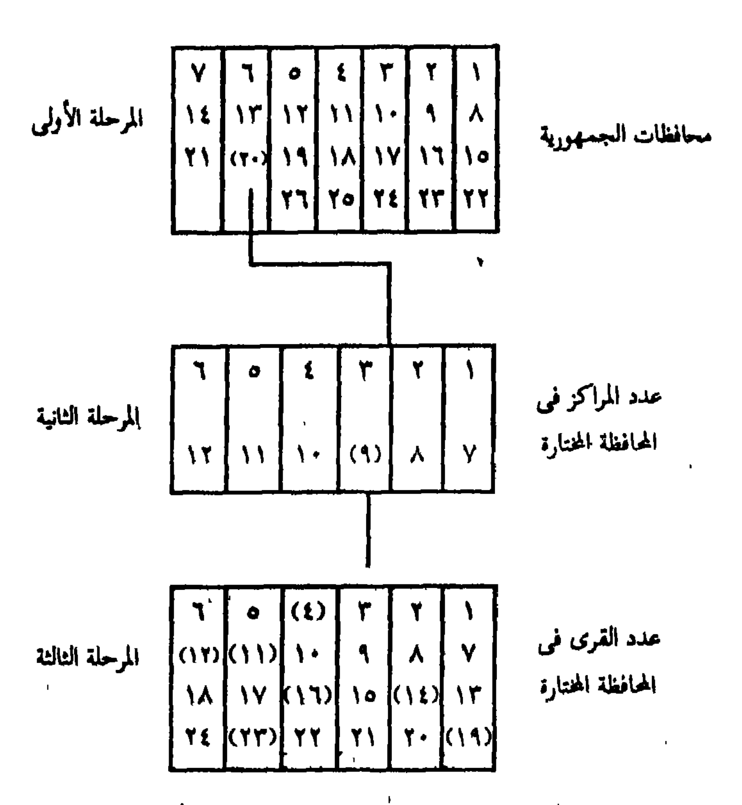
وتتطلب مثل هذه الدراسة حساب الخطأ المعيارى لفئات استخدام الأرض فى كل طبقة أو مستوى من مستويات الارتفاع وللعينة ككل. ويحسب الخطأ المعيارى لمثل هذه البيانات (بيانات نسبية) على أساس أنه يساوى الجزء التربيعى لحاصل ضرب نسبة وجود الظاهرة (أى فئة من فئات استخدام الأرض) فى نسبة عدم وجودها فى حجم العينة. ولنأخذ فئة الاستخدام الثانية (أراضي الحشائش) لتوضيح ذلك كما يلى:

مستويات الارتفاع (متر)	تكرار أراضي الحشائش	حجم العينة ن	أ – نسبة وجود أراضي الحشائش في العينة	ب – نسبة عدم وجود أراضي الحشائش في العينة	الخطأ المعيارى للنسب = $\sqrt{أ \times ب \times ن}$
أقل من ١٥٠	٣	٤	١	–	٨
١٥٠ – ٣٠٠	٥	٢١	٥	٣٠	٤١
أكثر من ٣٠٠	–	٦	–	٤٥	٥١
المجموع	٨	٣١	٦	٥٥	١٠٠

ويمكن بعد ذلك تحويل الخطأ المعيارى فى كل مستوى (طبقة) من مستويات الارتفاع وللعينة كلها إلى نسبة مئوية ($\frac{\text{الخطأ المعيارى}}{\text{حجم العينة}}$ × ١٠٠). فيمكن مثلاً التعبير عن التكرار المطلق لمواقع العينة الفرعية لمستوى الارتفاع ١٥٠ - ٣٠٠ متراً (٢١ موقعاً من جملة المواقع لهذا المستوى وهى ٤١ موقعاً وبخطأ معيارى ٩,٣٧٢) بتكرارات أخرى نسبية، كأن نقول أن التكرار النسبى الحقيقى لمواقع هذه الفئة (أراضى الحشائش) من فئات استخدام الأرض فى المنطقة فى مستوى الارتفاع ١٥٠ - ٣٠٠ متراً بدرجة ثقة أو عند المستوى الاحتمالي ٩٥٪ يمكن أن ينحصر بين ٥,٤٨٪ و ٩٦,٩٢٪.

د- العينة العشوائية المتعددة المراحل Multi-Stage Random Sample

يلائم هذا النوع من العينات العشوائية دراسة المجتمعات الكبيرة، مثل الدراسات السكانية أو الدراسات فى مجال الجغرافية الاقتصادية، وهى مجتمعات يمكن تقسيمها إلى عدد من الأقسام المتشابهة التى يحتوى كل قسم منها على عدد من المفردات التى تتصف بعدم التجانس فى خصائصها، وإذا أطلق على هذا النوع من العينات بأنه «متعدد المراحل». فمثلاً إذا أردنا دراسة الحالة الاجتماعية فى الريف على مستوى محافظات الجمهورية، فإننا نقوم أولاً باختيار عشوائى لعدد من محافظات الجمهورية، وبعد ذلك فى مرحلة تالية تقوم باختيار عشوائى لعدد من مراكز المحافظات المختارة سابقاً، ثم تأتى بعد ذلك المرحلة الثالثة وفيها نقوم باختيار عشوائى لعدد من قرى المراكز المختارة فى المرحلة الثانية، وتكون هذه القرى بما يختار منها عشوائياً من أسر عبارة عن المفردات التى تجرى عليها لتحديد بعض المؤشرات والمقاييس الاحصائية. ويهدف التدرج السابق فى أخذ العينات فى مراحل إلى التبسيط والمحافظة على طبيعة المفردات غير المتجانسة داخل العينة فى كل مرحلة من المراحل. ويوضح ذلك الشكل التالى.



ويعاب على المعاينة المتعددة المراحل أن كثرة عدد المراحل التى قد تتضمنها تضعف العلاقة بين معالم المجتمع الأصلي وخصائص العينة مما يؤثر بالتالى على تقدير معالم المجتمع من بيانات العينة المتحصل عليها فى آخر مرحلة، كما أن هذا النوع من العينات يتطلب من الباحث جهداً كبيراً فى تحديد حدود أو إطار كل مرحلة وتحديد حجم العينة الفرعية المطلوبة من كل منها وذلك فى ضوء الاعتبارات الخاصة بالاختيار فى المعاينة العشوائية.

ثانياً: العينات غير العشوائية (العمدية) Non-Random Samples

كثيراً ما يتعرض أسلوب المعاينة العشوائية لبعض العقبات التى تحول دون التمسك به أو الاعتماد عليه في دراسة المجتمعات، وذلك عندما يتطلب سحب العينة العشوائية امكانيات مادية وفنية، أو عندما يجد الباحث صعوبة فى الوصول إلى وحدة من وحدات المجتمع المختارة، أو فى حالة عدم معرفة كل مفردات المجتمع الذى ستسحب منه العينة العشوائية. وفى مثل هذه الحالات يضطر الباحث إلى

Converted by Tiff Combine - (no stamps are applied by registered version)

اتباع أسلوب التعمد والتحيز الشخصي في اختيار مفردات العينة، أو ما يعرف بأسلوب العينة العمدية (غير الاحتمالية). وبذلك يقوم اختيار هذا النوع من العينة على أساس شخصي ولاتراعي فيه الفرص المتكافئة للمفردات لأن تكون ضمن العينة، أي لاتراعي فيه صفة العشوائية.

وكثيراً ما تستخدم المعاينة العمدية، بصفة عامة، في الأبحاث الاستطلاعية. كما في حالة تقدير معالم مجتمع كبير، أو عند محاولة معرفة فكرة تقريبية سريعة عن خصائص ظاهرة ما بحيث لاتستخدم نتائجها للتعميم علي المجتمع. كما تستخدم المعاينة العمدية في الاختبارات القبلية (السابقة) Prior Tests مثل اختبار صحيفة الأسئلة لمعرفة مدى تجاوب المبحوثين حتى يمكن إجراء التعديلات اللازمة في الأسئلة قبل بدء المعاينة الرئيسية، أو في حالة القيام ببعض القياسات لظاهرة ما لمعايرة الأجهزة المستخدمة في القياس والتأكد من سلامتها. وجدير بالذكر أنه لاتوجد هناك أية طريقة احصائية لمعرفة وقياس دقة نتائج المعاينة العمدية غير الاحتمالية، ولذا لاتعد هذه الطريقة من طرق المعاينة الجيدة، إلا أنه في بعض الأحيان قد لايجد الباحث سبيلاً ممكناً عملياً للمعاينة سوى استخدام هذه الطريقة. وسنعرض فيما يلي لأهم أنواع العينات غير العشوائية (العمدية أو غير الاحتمالية) وهي العينة الغرضية، العينة بالحصة، والعينة العنقودية.

أ- العينة الغرضية:

تلائم طريقة العينة الغرضية الدراسات التي تخص الظاهرات التي تشتد فيها درجة تباين متغايراتها، مما يجعل الباحث مضطر إلى تحديد واختيار المتغيرات الخاصة بالبيانات المراد جمعها والتي يرى من وجهة نظره أنها تصلح للدراسة. فمثلاً الباحث الذي يدرس مستوى المعيشة في الريف المصري لايمكنه الاعتماد على الاختيار العشوائي لعينة من القرى، بل يعتمد على تحديد عدد من القرى تمثل في نظره مجتمع القرى المصرية وتكون محلاً للدراسة.

ب- العينة بالحصة Quata Sample:

يضطر الباحث إلى استخدام مثل هذا النوع من العينات العمدية (غير الاحتمالية) عندما يتطلب منه القيام بإجراء عدد معين من المقابلات لأشخاص لهم صفات محددة فى مكان معلوم، أو بإجراء عدد معين من الزيارات الميدانية لجمع بيانات عن ظاهرة معينة داخل منطقة محدودة. وفى طريقة العينة بالحصة لاتختار مفردات (وحدات) العينة عشوائياً ولكن تستخدم أية معلومات تساعد على الحصول على الحصة المطلوبة بسرعة وبتكاليف قليلة. ولذلك فإن هذه الطريقة تستخدم بكثرة فى معاينة واستطلاع الرأى العام كما هو متبع فى معهد جالوب بالولايات المتحدة الأمريكية عند التنبؤ بنتيجة الانتخابات العامة، إذ يطلب من الباحثين فى هذه الحالة التعرف على رأى مجموعة من الناخبين على أن تكون من بينهم نسبة معينة من فئات مختلفة مثل فئات أصحاب المهن الحرة، وفئة العمال وفئة الموظفين ... إلخ، ويترك للباحثين حرية التصرف فى اختيار الأعداد المطلوبة منهم بأية طريقة يجدونها سهلة ومناسبة.

وقبل إجراء العينة بالحصة يجب التأكد من مجموعة الخصائص (ثلاث أو أربع خصائص مثلاً) التى تميز المجتمع الأصلى بحيث ترتبط هذه الخصائص ارتباطاً وثيقاً بالمتغير قيد البحث، وتصمم عينة تكون ممثلة لهذه الخصائص مجتمعة. ويتضمن تصميم العينة بالحصة ثلاث مراحل هى:

١ - مرحلة تصنيف المجتمع الأصلى على أساس الخصائص موضع الدراسة.

٢- مرحلة تحديد نسبة المجتمع فى كل طبقة أو فئة.

٣- مرحلة تحديد الحصص التى يراد دراستها واختيارها فى ضوء العدد المطلوب.

وجدير بالذكر أنه يمكن اعتبار العينة بالحصة نوع من العينات الطبقية التى يكون فيها الاختيار داخل الطبقة إختيار غير عشوائى، مما قد يؤدى إلى الوقوع فى خطأ التحيز من قبل الباحث من جراد التصنيف الشخصى للعناصر والفئات من ناحية وعدم عشوائية الاختيار من ناحية أخرى.

حـ- العينة العنقودية Cluster Sample:

تشبه طريقة المعاينة العنقودية للعينة متعددة المراحل فى كثير من مراحل إجرائها. وتقوم هذه الطريقة على أساس إختيار مفردات العينة فى حزم أو عناقيد Clusters بأقل جهد وتكلفة مما يحدث بالنسبة للعينة العشوائية. فمثلاً إذا كنا بصدد دراسة مستوى المعيشة فى منطقة متخلفة بأحد الأقسام الإدارية فى محافظة الإسكندرية، وبإفتراض عدم وجود سجل يضم سكان هذه المنطقة، ولكنهم يوجدون فى سجلات مصلحة الكهرباء للمحافظة كلها بما فيها المنطقة قيد البحث، فإنه يمكن اختيار العينة على عدة مراحل أو تدريجياً بشرط أن تكون متكاملة.

ولإجراء سحب عينة عنقودية، نفترض أن دراسة جغرافية اجتماعية (مثل دراسة خصائص النشاط السكاني والعوامل المؤثرة فيه) ستجرى على سكان مدينة ما يبلغ تعدادها ٣٠٠٠٠ نسمة والمسجلين فى قوائم يمكن الحصول عليها بسهولة. والمطلوب أن تكون العينة التى تجرى عليها لدراسة مكونة من ٣٠٠٠ شخص فقط. ففى هذه الحالة تحتم طريقة العينة العنقودية أن تكون العينة متركزة (أو متجمعة) فى أجزاء قليلة من المدينة. فإذا افترضنا أن المدينة تنقسم إلى ٥٠ شياخة فى كل منها ٦٠٠ شخصاً فإنه يمكن اختيار عينة من خمس شياخات فقط (أى شياخة واحدة لكل ١٠ شياخات) وتجرى لدراسة على هذه الشياخات الخمس بما تحتوى من سكان.

من العرض السابق لكل من أسلوبى الحصر الشامل والمعاينة يتضح أن أسلوب الحصر الشامل يتعرض لخطأ التحيز الذى ينتج عن أى تقصير أو إهمال فى خطوات البحث، بينما يتعرض أسلوب المعاينة لخطأ الصدفة أو الخطأ العشوائى، وهو خطأ يصاحب عملية الاختيار نفسها. كما يتعرض أسلوب المعاينة لخطأ التحيز إذا حدث أى خطأ أو تقصير فى خطوات البحث بالعينة. ولكن تفضل الدراسة بالعينة كثيراً على الدراسة بالحصر الشامل لأن العينة كجزء من المجتمع تعطى فرصة كبيرة للتحكم والسيطرة على جميع العوامل المحيطة أو الخارجية. كما يمكن استخدام الخطأ العشوائى للعينة فى تقدير مستوى الثقة التى تعمم بها النتائج على المجتمع من بيانات هذه العينة.

طرق (أدوات) جمع البيانات:

بعد أن يتم تحديد الأسلوب الذى على أساسه سيتم جمع البيانات سواء كان أسلوب الحصر الشامل أو أسلوب العينات، فإنه يجب اتباع إحدى الطرق أو الوسائل (الأدوات) الاحصائية التى تستخدم فى عملية الحصول على البيانات لإتمام العمل الاحصائى لأى من الأسلوبين. وأهم طرق أو أدوات جمع البيانات من مصادرها: طرق المراسلة والاتصال والعمل الحقلى (الميدانى)، وتنقسم كل أداة منها إلى مجموعة من الوسائل يتم بواسطتها إتصال الباحث بمفردات المجتمع أو العينة. ولكل أداة مزايا متعددة فى استخدامها بما يتفق مع طبيعة البحث. فالمراسلة والاتصال أكثر الأدوات كفاءة فى التعداد والبحوث والدراسات السكانية، والعمل الحقلى أو المسح الميدانى بما يتضمن من وسيلة الملاحظة أو المشاهدة أكثرها كفاءة فى دراسة سلوك الظواهر الجغرافية الاجتماعية أو فى الكشف عن تفاصيل الظواهر وعن الصلات الخفية التى توجد بين عناصرها أو بينها وبين بعض الظواهر الأخرى.

وجدير بالذكر أن اختيار أى أداة من أدوات جمع البيانات يتوقف على طبيعة المعلومات التى يراد جمعها والوقت المسموح به والامكانيات المادية المتاحة للباحث.

أولاً: المراسلة والاتصال:

يعتمد الباحث على هذه الأداة فى جمع البيانات إذا تعذر الوصول أو الاتصال المباشر بمفردات المجتمع أو العينة. وتتم عملية جمع البيانات بهذه الأداة عن طريق إرسال استمارة البيانات الاحصائية بالبريد، أو بواسطة الاتصال التليفونى بالمبحوثين، أو حتى عن طريق نشر الأسئلة على صفحات المجلات أو الجرائد المتخصصة.

أ- **المراسلة بالبريد:** يقوم الباحث فى هذه الطريقة بإرسال رسالة للشخص الذى وقع عليه الاختيار لاستبيانه يوضح له فيها أهمية البحث وأهدافه وسرية البيانات وعدم استخدامها إلا لغرض البحث فقط، ومع الرسالة يرفق الباحث الاستمارة الاحصائية المطلوب الإجابة على أسئلتها كما يرسل مع الرسالة مظروف

خاص بعنوان الباحث وخالص الرسوم البريدية، ويطلب من المبحوث إعادة الاستمارة الإحصائية مرة أخرى إلى الباحث.

ومن يرى هذه الطريقة أنها سهلة التنفيذ وقليلة التكاليف، إلى جانب أنها تعطى فرصة كافية للمبحوث فى التفكير والإجابة على الأسئلة دون ما حرج أو تردد. هذا بالإضافة إلى أنها تجنب الباحث خطأ التحيز الذى قد يظهر فى طريقة العمل الحقلى والاتصال المباشر بمفردات العينة أو المجتمع. وعلى الرغم من ذلك يعاب على طريقة المراسلة بالبريد أن نسبة الإجابات تكون عادة قليلة، وبصفة خاصة إذا كانت الاستمارة الإحصائية تحتوى على عدد كبير من الأسئلة فإن ذلك يكون سبباً فى إهمال المبحوثين للاستمارات وعدم استيفائها وإعادتها للباحث. كما تتضح صعوبة هذه الطريقة فى حالة إذا كان عدد من مفردات العينة يجهلون القراءة والكتابة. هذا بالإضافة إلى أن هذه الطريقة تحتاج إلى دقة بالغة فى وضع الأسئلة، إذ أنه قد ينتج عن عدم فهم المبحوثين بعض الأسئلة وقوعهم فى أخطاء تؤثر على دقة النتائج مثل خطأ التحيز فى الإدلاء بالمعلومات لانعدام الرقابة على الإجابات ولاعتقاد المبحوثين بعدم جدية أو ضرورة الدراسة.

ب- **الاتصال التليفونى:** تصلح طريقة الاتصال التليفونى، كطريقة من طرق جمع البيانات، فى الدراسات المحدودة التى يلعب فيها عامل الوقت دوراً مؤثراً، والتى يضمن فيها الباحث وجود أجهزة تليفونية لدى المبحوثين الذين تتكون منهم العينة أو المجتمع. ويمكن بهذه الطريقة مخاطبة المبجوثين والحصول منهم على الإجابات للأسئلة الموجهة إليهم.

وعلى الرغم من أن الإتصال التليفونى يعد من أسهل الطرق لجمع البيانات إلا أنها أصعبها من حيث إمكانية الحصول على نسبة عالية من الإجابات إذا كانت الأسئلة طويلة وتحتاج لوقت طويل فى فهمها، لذا فلابد أن تكون الأسئلة قصيرة وسريعة حتى لاتأخذ وقتاً طويلاً فى الإجابة عليها.

ثانياً: العمل الحقلى (الميدانى) Fieldwork

العمل الحقلى أو الميدانى كما عرفه Wooldige and East (1967) هو عبارة عن «الفحص القريب والتحليل فى الميدان لجزء من البلاد، بما فيه من ظواهر

طبيعية وبشرية، تكون سهلة الوصول، وموضحاً مظهر أو أكثر من مظاهر الاختلاف المكاني، وبذلك يتميز العمل الحقلي بأنه يضع الباحث وجهاً لوجه أمام مفردات ومتغيرات الظاهرة أو الظواهر المراد دراستها وتحليلها، كما أن الباحث في الميدان يستطيع أن يرى ويلمس الجوانب غير الواضحة عن الظاهرة وبالتالي يتأكد من صحة البيانات والمعلومات السابقة عنها. ولذا فإن نجاح الباحث في دراسته يتوقف إلى حد كبير على نوعية وكيفية العمل الحقلي الذي أجراه، وعلى الوقت والجهد الذي بذله، وعلى الزيارات التي قضاها في منطقة البحث.

ويشمل العمل الحقلي طريقة: المقابلة الشخصية (الاتصال المباشر) للمبحوثين، أو الملاحظة الميدانية وقياس الظواهر في الطبيعة، أو المسح الميداني والزيارات للمزارع والمصانع. ويتوقف استخدام كل طريقة منها على خطة البحث ونوع الدراسة، كما أن لكل منها مزايا وعيوب نوضحها فيما يلي:

أ- المقابلة الشخصية Intervewing:

تعرف المقابلة أحياناً بطريقة الاتصال المباشر لجمع البيانات، إذ يتم فيها انتقال الباحث إلى المبحوثين (مفردات العينة) وذلك بغرض المواجهة الشخصية للحصول على المعلومات التي تحتاجها الدراسة كما في حالة دراسة الخصائص الاجتماعية والثقافية لسكان أحد الأقسام الإدارية في محافظة ما.

وفي حالة دراسة مفردات مجتمع يتميز بأن عدد مفرداته كبير يستعين الباحث بمندوبين لجمع البيانات يشترط فيهم أن يكونوا مدربين تدريباً كافياً على العمل بهذه الطريقة، ويتصفون بالأمانة في تدوين البيانات.

ويمكن أن تتم المقابلة إما في أشكال محددة، أو في صورة غير محددة. فهناك المقابلة المحددة أو المقفولة Closed Interview، وهي المقابلة المقننة أو المنهجية، التي تتخذ أسلوباً منظماً حيث تكون حالة وضع الأسئلة سابقة على المقابلة نفسها، وتوجه نفس الأسئلة لجميع المفردات بدون تغير سواء في الأسلوب أو الصياغة. وهناك أيضاً المقابلة غير المحددة أو المفتوحة، وهي المقابلة غير المقننة أو غير المنهجية ،

التى تتميز بالأسئلة الحرة التى تتواتر بطريقة طبيعية تلقائية، أى لاتلتزم باستخدام صياغة الأسئلة وأسلوبها.

ومن أهم مزايا المقابلة الشخصية أنها تلائم كثيراً دراسة المناطق التى ترتفع نسبة الأمية بين سكانها، كما تتيح للباحث الحصول على معلومات أولية تقل فيها الأخطاء الصادرة من المبحوثين إلى درجة كبيرة. كذلك تعطى هذه الطريقة الفرصة لتوضيح الأسئلة الغامضة أو التى تبدو غير مفهومة للمبحوثين، هذا بالإضافة إلى أنه يمكن للباحث إضافة بعض الأسئلة التى يرى إضافتها، أو حذف البعض الآخر تبعاً لما تمليه ظروف المقابلة، كما أن الباحث يستطيع كشف أى تناقض يمكن أن يقع فيه المبحوثين.

أما عيوب طريقة المقابلة فمن أهمها إحتمال تحيز الباحث، أو توجيهه لمفردات المبحوثين لوجهة نظر لاتخدم غرض البحث مما يؤثر على دقة النتائج، أو قد تتضمن هذه الطريقة بعض الغش إذ من المحتمل أن يقوم الباحث باستكمال بعض الإجابات بنفسه ليتسنى له إتمام عمله فى وقت قصير. وتحتاج طريقة المقابلة جهود مضنية واعتمادات مالية كبيرة إذ أنها تحتاج فى بعض الأحيان إلى عدد كبير من المندوبين لجمع البيانات. كما أن هذه الطريقة لاتتمشى مع الدراسات التى تتصف بأنها تأخذ طابعاً خاصاً، بمعنى أن تكون أسئلتها محرجة أو حساسة للأفراد الذين يجرى عليهم الباحث والاستقصاء والذين ربما يحجمون عن الإجابة على مثل هذا النوع من الأسئلة.

ب- الملاحظة الميدانية وقياس الظواهر فى الطبيعة:

تعرف الملاحظة الميدانية بأنها المشاهدة الدقيقة لظاهرة ما، مع الاستعانة بأساليب البحث والدراسة لقياس وتسجيل كافة أوجه التغيرات Variations (مكانية أو زمنية) فى الظاهرة وفق خطة معينة تتلائم مع طبيعة تلك الظاهرة. وهى بذلك تستخدم فى الدراسات والأبحاث التى تهدف إلى اختبار الأسس والنظريات Basis and theoris التى تضبط ظاهرة أو مشكلة معينة. فعند دراسة ظاهرة أو مشكلة ما، مثلاً، فإننا نضع الفروض المناسبة لدراستها وحلها على أساس وضع خطة محددة

تشتمل على بعض التجارب العلمية أو القياسات الحقلية، باستعمال بعض الأجهزة لقياس وتسجيل المتغيرات المتعلقة بالفروض الموضوعة، ثم نقوم باختبار صحة الفرض واحداً بعد الآخر مع استبعاد الفرض الذى لاتثبت صحته وأهميته.

وعند إجراء عملية الملاحظة فإنه يجب على الباحث مراعاة الحرص والدقة فى القياس وعدم التحيز. فمثلاً عند تحليل الاختلافات المكانية أو الزمنية فى شكل شاطئ منطقة ساحلية فإنه يجب قياس وتسجيل المتغيرات التى تؤثر فى هذه الاختلافات حتى يمكن الوقوف على الطريقة التى يتأثر بها الشاطئ، ولرسم وتحديد المناطق التى تحدث بها هذه الاختلافات أكثر من غيرها من المناطق. ومن المتغيرات التى يمكن قياسها لهذا الغرض: انحدار الشاطئ Beach Slope حجم الرواسب، خصائص الأمواج (طول الموجة، ارتفاعها، انحدارها وقوتها، زمن الموجة واتجاهها)، وخصائص التيارات الساحلية من حيث اتجاه التيار وقوته، والعوامل الجوية من ضغط جوى واتجاه الرياح وقوتها.

ويسود استخدام الملاحظة أو المشاهدة كأداة من أدوات البحث لجمع وتسجيل المعلومات فى مجال الدراسات المعملية أو فى التجارب الحقلية، وذلك بقصد تحقيق فرض معين، أو لمعرفة علاقة بين متغيرين، أو لإيضاح بعض النتائج التجريبية التى يكون معناها مازال غامضاً أو مبهماً. وأثناء الملاحظة يقوم الباحث بتسجيل الملاحظات أو المشاهدات فى شكل قياسات معملية أو حقلية عن المتغيرات التى تحكم التجربة مثل رصد الحركة والوقت، كما فى حالة رصد حركة المرور على أحد الطرق فى فترات زمنية محددة. وتجدر الإشارة هنا إلى أن الملاحظة أو المشاهدة عن طريق التجربة تمكن الباحث من السيطرة على بعض المتغيرات المحيطة بالظاهرة قيد البحث والتى لاتدخل ضمن المتغيرات التى يراد دراساتها حتى يمكن اختبار سلوك الظاهرة بدقة، والوقوف على حقيقة التأثير النسبى لمختلف المتغيرات. ونظراً لصعوبة التحكم فى جميع المحددات والعوامل الجغرافية، وهى عوامل متشابكة، فإن الملاحظة أو المشاهدات عن طريق التجربة الميدانية فى الدراسات الجغرافية لاتتم إلا إذا كان مجال الظاهرة محدوداً والمتغيرات قيد البحث عددها قليلاً.

جـ- الزيارات:

يشمل العمل الحقلى (الميدانى) أيضاً جمع البيانات المنشورة التى تفيد دراسة وتحليل الظاهرة قيد البحث كالتقارير والوثائق والاحصاءات من الشركات والمؤسسات أو الهيئات الخاصة والحكومية فى منطقة البحث. وكما يقول Wool-drige and East (1967) أن «العمل الحقلى لايعد عملاً حقيقياً إلا إذا شمل الزيارات للمزارع والمصانع ومراكز الاحصاء». ومما تجدر الإشارة إليه فى هذا الصدد أنه يجب مطابقة النشرات والتقارير التى يجمعها الباحث من مصادرها على الطبيعة للوثوق من سلامتها العلمية، وللتأكد من صحة ما تحتويه من بيانات، كما يجب التعرف على الطرق والأساليب التى جمعت بواسطتها هذه البيانات ويتم ذلك عن طريق مناقشة المختصين وذوى الخبرة فى الهيئات المسئولة عن نشر البابنات. وإذا ما تعذر الحصول على البيانات المطلوبة أثناء القيام بالزيارات الميدانية فإنه لابد أن يقوم الباحث بتصميم استمارة احصائية، أو ما يعرف «بالاستبيان» لاستكمال هذا النقض بنفسه عن طريق الاستفسار الشخصى وتدوين المعلومات عن كل أو بعض المتغيرات المطلوب دراستها.

من العجالة السابقة عن العمل الحقلى نلاحظ أن إجراءاته ووسائله تتنوع بتنوع الظروف والعوامل المتحكمة فى الظاهرة قيد البحث، ولكن قد يستعين الباحث فى الميدان ببعض الوسائل التى تعينه فى الدراسة الميدانية بصفة عامة وفى الملاحظة بصفة خاصة وهى:

(١) تسجيل القياسات التى أجريت على المتغيرات المطلوب دراستها بواسطة الأجهزة والمعينات الخاصة، وتدوين المشاهدات الميدانية إما فى جداول وإما على أجهزة تسجيل إذا تيسر استعمالها. ولتدوين القياسات والمشاهدات مكانة خاصة فى الدراسات المنظمة الهادفة لأهميتها فى تكوين العلاقات بين المشاهدات والملاحظات المختلفة.

(٢) الخرائط: تعتبر الخريطة من أصلح الوسائل لمعرفة العلاقات المكانية وتفسير وتعليل كثير من غوامض الظاهرة أو الظاهرات المدروسة. ونظراً لأنه يستحيل

على الباحث أن يزور كل مكان ويفحصه في الطبيعة، فلابد له من الاعتماد على الخريطة لمعرفة الأماكن التي يصعب عليه رؤيتها في الطبيعة. وعموماً فإن الباحث عند قيامه بالزيارات أو ملاحظة وقياس متغيرات ظاهرة ما في الطبيعة يحتاج إلى عدة خرائط ضرورة عن الظروف الطبيعية لمنطقة الدراسة من أهمها الخريطة الكنتورية والطبوغرافية، والخريطة المناخية، وخريطة استخدام الأرض في الأغراض المختلفة، وخريطة السكان. ومما يدل على أن الخريطة هي أداة الجغرافي الأولى في العمل الحقلي ما قاله H.R. Mill في ذلك «إن ما لايمكن إثباته على خريطة لايمكن وصفه» (Wooldrige and East, 1967). وبناء على ذلك أصبحت الخرائط ذات أهمية خاصة في مختلف فروع العلم والمعرفة.

(٣) الصور الجوية والفوتوغرافية لظاهرات ومواقف معينة: وهما من الوسائل المعينة في الدراسة الميدانية. فمن المعروف أن الباحث قد يرى الموقف من خلال طريقة تفكيره الخاصة وإهتماماته الشخصية بالدراسة التي يقوم بها، ولكن لايتفق باحثان على وصف واحد لظاهرة معينة أو لموقف معين إذ أن كلا منهما يراه من وجهة نظره الخاصة.

وعموماً إذا كان منهج البحث هو الذى يحدد الطريقة المتبعة فيه، فإن الطريقة - بالتالي - هي التي تحدد أداة أو وسيلة جمع البيانات الأكثر مناسبة. إلا أن هذا لايعني الاعتماد على أداة واحدة فقط من الأدوات السابق ذكرها لجميع البيانات، لأنه يمكن جمع المعلومات بأكثر من أداة تبعاً لما تمليه طبيعة الدراسة ونوعية المعلومات المطلوب الحصول عليها. ونظراً لأن هذه الأدوات غير مستقلة تماماً عن بعضها فإن الباحث يستطيع أن يحدد الأدوات التي سيستخدمها في جمعه للبيانات والمعلومات الدقيقة التي يتعرض لها بالتحليل الاحصائي بحيث يحصل في النهاية على نتائج يتخذ على أساسها القرارات فيما بعد.

الاستمارات الاحصائية:

بعد أن يتم اختيار طريقة جمع البيانات والمعلومات عن خصائص مفردات

المجتمع أو العينة، سواء بأسلوب الحصر الشامل أو المعاينة (العينات)، يلجأ الباحث إلى عمل استمارة إحصائية خاصة، تتضمن أسئلة محددة عن تلك الخصائص المراد معرفتها وقياسها، ولتكون مرشداً له فى جمع بياناته ورسم إطاراً محدداً لها، هذا فضلاً عن استخدامها كأداة لتسجيل البيانات أو قناة تستقى المعلومات من خلالها. وعادة ما تستخدم الاستمارة الإحصائية فى الدراسات التى تحتاج إلى جمع بيانات كثيرة قابلة القياس ويمكن تسجيلها بانتظام.

وكما ذكرنا آنفاً أن خطوات تصميم البحث تبدأ بوضع إطار للبيانات التى يجب الحصول عليها لاستخدامها فى حل مشكلة البحث، ثم تحديد مصادر هذه البيانات والوسائل التى ستتبع فى الحصول على هذه البيانات. وفى المرحلة الأخيرة ذكرنا أنه يجب أن تختار وسائل جمع البيانات إختياراً سليماً يبنى على أساس مدى ملائمة كل وسيلة، من حيث مزاياها وعيوبها، لأهداف البحث، وعادة ما تكون الاستمارة الإحصائية أداة هامة من أدوات أو وسائل جمع البيانات وتستخدم بالإضافة إلى الأدوات أو الوسائل الأخرى مثل المقابلة أو الملاحظة.

وهناك نوعين رئيسيين من الاستمارات الاحصائية هما: كشف البحث Schedule، وصحيفة الاستبيان Questionnaire، ولكل منهما مزاياه وعيوبه التى نوضحها فيما يلى:

١- **كشف البحث**: يطلق إسم كشف البحث على الاستمارة الاحصائية التى تضم مجموعة من الأسئلة التى تسأل وتدون بواسطة الباحث فى مقابلة شخصية للبحوث (مفردة البحث) الذى وقع عليه الاختيار فى عينة البحث. كما يضم كشف البحث عند استخدامه فى الملاحظة أو المشاهدة الميدانية بيان بمتغيرات الظاهرة المطلوب قياسها وجمع المعلومات عنها. ويتيح هذا النوع من الاستمارات الاحصائية للباحث درجة عالية من المرونة عند تصميمها بإعطاء الفرصة فى إضافة أو حذف ما يراه الباحث من أسئلة تبعاً لظروف المقابلة الشخصية، أو صرف النظر عن بعض العوامل التى قد لايكون لها دوراً يذكر فى تباين الظاهرة موضع الدراسة. إلا أنه من أخطر عيوب كشف البحث هو إحتمال تحيز الباحث لوجهة نظر شخصية لاتخدم غرض البحث، مما يؤثر على دقة النتائج التى ينتهى إليها البحث.

٢- صحيفة الاستبيان: وهى عبارة عن الأداة التى تستخدم للحصول على البيانات عن طريق الإجابة على أسئلة تتعلق بالظاهرة قيد البحث والتى يجيب عليها المبحوث بنفسه، وهذه قد ترسل بالبريد أو تسلم باليد للمبحوث الذى يطلب منه فى كلتا الحالتين إعادتها للباحث بعد إستيفائها.

وتجدر الإشارة هنا إلى توضيح الفرق بين الاستبيان وصحيفة الاستبيان حتى لايختلط الأمر بينهما. فالأولى عبارة عن وسيلة قائمة بذاتها لجمع المعلومات بطريقة سريعة عن موضوعات محددة ومن مجموعة كبيرة من المفردات (المبحوثين)، بينما تستخدم الثانية كأداة لهذه الوسيلة التى يكون هدفها الأساسى ترجمة البحث العلمى إلى أسئلة معينة. وبصفة عامة تتميز صحيفة الاستبيان بسهولة تنفيذها وتوفيرها للوقت والتكاليف المادية، وإتاحتها الفرصة للمبحوث فى التفكير والإجابة على الأسئلة الحرجة دون تردد، بالإضافة إلى أنها تجنب للباحث الوقوع فى خطأ التحيز لعدم إمكانية فرضه لرأى معين أو لوجهة نظر خاصة. إلا أن إمكانية وجود أخطاء ناجمة عن تحيز المبحوث نفسه فى إجابة الأسئلة يعتبر من أهم مثالب صحيفة الاستبيان، بالإضافة إلى أنها لاتصلح تماماً إذا كانت مجموعة المبحوثين فى العينة أو المجتمع تحتوى على عدد كبير يجهل القراءة والكتابة، أو إذا كانت البيانات المطلوبة كثيرة ووقت المبحوث ضيقاً مما يؤدى إلى تكاسل المبحوث فى استيفاء الاستمارة وإعادتها للباحث.

تصميم الاستمارة الاحصائية:

مهما كانت طبيعة البيانات المطلوب الحصول عليها أو الوسيلة المتبعة فى جميع هذه البيانات فإنه يجب على الباحث مراعاة بعض الشروط الهامة عند تصميمه للاستمارة الإحصائية، لأن التصميم الجيد والصياغة المتقنة لأسئلة الاستمارة يعد أحد العوامل الجوهرية فى انجاح العمل الحقلى بصفة خاصة والبحث الذى يقوم عليه بصفة عامة. وتحتاج عملية تصميم الاستمارة الاحصائية من الباحث المعرفة الكاملة والدراية التامة بأصول صياغة الأسئلة. ورغم أن الاستمارات تختلف فى تصميمها، إلا أن هناك قواعد وشروط يجب توافرها حتى

يأخذ تصميم الاستمارة دورة في انجاح البحث. هذه الشروط منها ما هو متعلق بشكل الاستمارة، ومنها ما هو متعلق بمضمونها من حيث نوعية الأسئلة وطريقة وضعها وصياغتها.

١- **شكل الاستمارة:** لاشك أن الإهتمام بشكل الاستمارة الإحصائية يعتبر من العوامل الرئيسية فى عملية جمع البيانات الدقيقة غير المشكوك فيها، حيث يشجع الشكل الجيد المبحوثين على الاستجابة لمحتواها. ويتحدد شكل الاستمارة الجيد بعدة عوامل منها:

أ- جودة الاستمارة، من حيث نوع الورق المستخدم الذى يجب أن يكون من النوع الذى يتحمل الاستخدام الكثير من تدوين المعلومات.

ب- حجم الاستمارة، من حيث عدد صفحات الاستمارة التى يجب أن لاتكون قليلة على حساب الأماكن الحالية المخصصة للإجابة، أو لاتكون كثيرة حتى لايكون ذلك سبباً فى إرهاق المبحوثين فى الإجابة على أسئلتها.

حـ- ترتيب وتنظيم الأسئلة داخل الاستمارة، إذ أن التسلسل والترتيب فى وضع الأسئلة (عن طريق إعطاء الأسئلة أرقاماً تدريجية، أو وضع الأسئلة فى شكل مجموعات أو تقسيمات متجانسة تترابط فيما بينها ترابطاً منهجياً، يمكن معه حصر المطلوب، بحيث تبدأ من الأسئلة البسيطة إلى الأسئلة المركبة، أو من أسئلة عامة تتميز بالشمول، إلى أسئلة خاصة تتميز بالتركيز على أفكار دقيقة ومحددة). يعتبر من أهم الشروط التى يجب مراعاتها عند تصميم الاستمارة الاحصائية مهما كان نوعها، لأن ذلك يساعد على سهولة الإجابة، كما يعمل على تسهيل عملية التحليل والدراسة بعد ذلك.

وعموماً يجب أن يظهر عنوان البحث بوضوح فى صدر الاستمارة، وكذلك اسم الهيئة أو الجهة المشرفة على الدراسة، بالإضافة إلى ما يشير إلى سرية استخدام بيانات الاستمارة إلا لغرض البحث فقط، مع وضع بعض التعليمات المختصرة والمبسطة لتوضيح أهداف الدراسة إن أمكن ذلك. ونظراً لأن معظم التحليلات

الاحصائية تقوم بها فى الوقت الحاضر أجهزة الحاسب الآلى فمن المستحسن أن تتضمن الاستمارة رموزاً Codes حتى تسهل مهمة نقلها وتفريغها على البطاقات الخاصة بالحاسب الآلى.

٢- مضمون الاستمارة: يقصد بمضمون الاستمارة هو كيفية صياغة الأسئلة والتى تعد ذات أهمية بالغة فى الحصول على إجابات صحيحة وبالتالى على معلومات دقيقة. وكلما كانت الأسئلة، أو التعبير عما هو مطلوب واضح دون ما صعوبة أو تعقيد لفظى أو سوء فهم كلما سهلت مهمة الباحث والمبحوث فى نفس الوقت. وبصفة عامة فإنه يمكن تحقيق ذلك بأن تكون الأسئلة على شكل حوار طبيعى تلقائى. أى ليس المقصود بها أن نتوصل إلى إجابات معينة، مع تجنب الأسئلة الطويلة التى تزيد من احتمالات سوء الفهم - كما يجب أن تكون الأسئلة محددة ودقيقة حتى نحصل على معلومات صحيحة، أى يجب أن يعطى كل سؤال فكرة واحدة واضحة عما نطلب السؤال عنه. فمثلاً يبدو السؤال: أين كان ميلادك؟ غامضاً، والأفضل منه يكون السؤال: فى أى قرية أو مدينة كان ميلادك؟ وهو يبدو أكثر تحديداً ووضوحاً من السؤال الأول. كما يجب أن تكون الأسئلة بعيدة تماماً عن الأسئلة الحرجة ذات الحساسية البالغة. ويستطيع الباحث التحايل على ذلك بصياغة أسئلة غير مباشرة، فمثلاً يمكن التعرف على مقدرة ودخل العامل بطرح الأسئلة التى تستفسر عن طبيعة العمل الذى يقوم به العامل. على أنه يجب أن يكون الباحث لبقا وذكياً عند وضع الأسئلة حتى لايضع أسئلة توحى بإجابات معينة، أو أسئلة افتراضية تكون الإجابة عليها غير مفيدة. فمثلاً يمكن طرح السؤال: ما مقدار الأجر الإضافى الذى ترغب أن تحصل عليه شهرياً حتى يتحسن مستوى معيشتك؟ بدلاً من السؤال: هل تكون راضياً لو ارتفع مرتبك الشهرى إلى ٦٠ جنيهاً؟. كذلك يجب وضع تفسيرات محددة للمصطلحات التى تكون مجالاً للشك من حيث الفهم، وتوضيحات دقيقة للتعريفات المستخدمة مثل تعريف الأسرة أو الدخل. كما يجب أن تصاغ الأسئلة إما لتوضيح الآراء أو

الاتجاهات Attiudes أو لتوضيح الحقائق مثل السن والمهنة أو الملكية الزراعية أو العقارية.

وقبل إتمام صياغة الاستمارة الاحصائية، ينبغى على الباحث أن يتفهم طبيعة المبحوثين موضع الدراسة وذلك عن طريق تصميم استمارة استطلاعية Piolt Ques-tionnaire توزع على عينة ذات عدد محدود من الأفراد ليست لهم علاقة بالبحث ليجيبوا على أسئلتها، ومن طريقة الإجابة فى الاستمارة الاستطلاعية يمكن التعرف على الأسئلة التى يمكن أن تكون غامضة أو غير مفهومة لتعاد صياغتها بعد توضيحها. كما تجدر الإشارة فى هذا الصدد إلى أنه يجب على الباحث أن يضع بعض الأسئلة للمراجعة Checking Questions للتأكد من صحة الإجابات، خصوصاً عند وجود تعارض فى الإجابات على هذا النوع من الأسئلة وإجابات الأسئلة الخاصة فى الاستمارة والتى تحمل نفس الإجابة أو عكسها. فمثلاً السؤال: هل تحب عملك؟ يتعارض مع السؤال: هل تتغيب كثيراً عن العمل؟. فإذا كانت الإجابة على السؤال الأول بالإيجاب وعلى الثانى بالنفى فإن ذلك يؤكد أن حب العمل لايؤدى إلى التغيب كثيراً عن العمل.

وفيما يلى مثال لاستمارة إحصائية عن دراسة العمران فى إحدى قرى منطقة زراعية مستصلحة حديثاً:

جامعة الإسكندرية
كلية الآداب
قسم الجغرافيا

استمارة بحث رقم (١)

دراسة العمران

(هذه الاستمارة سرية للغاية ولاتستخدم بياناتها إلا فى الأغراض العلمية)

أولاً: الحالة الاجتماعية للوافدين:

(١) عدد الأسر فى المسكن ( )

	(٢)	(٣)	(٤)	(٥)	(٦)	(٧)	(٨)
الاسم	النوع	السن	المهنة	الحالة الاجتماعية	المؤهل	محل الميلاد	الدخل الشهرى
١-							
٢-							
٣-							
٤-							
٥-							
٦-							
٧-							
٨-							
٩-							
١٠-							
١١-							
١٢-							
١٣-							
١٤-							
١٥-							

(٩) الجهة الوافد منها:

(١٠) سنة القدوم إلى القرية:

(١١) المهنة عند القدوم:

(١٢) هل تسافر إلى موطنك الأصلي:

ثانياً: السكن:

ب- ملكية السكن: (١) ملك خـاص (٢) ملك أهالى (٣) ملك حكومى

(٤) الإيجار الشهرى (٥) المساحة التى يشغلها المبنى ؟

جـ- ارتفاع المبنى: (١) دور واحد (٢) دورين (٣) أكثر من دورين (٤) عدد الغرف

(٥) هل توجد حظيرة الحيوان بداخله أم توجد خارجه

د- المرافق الصحية: (١) داخل المسكن.

(٢) خارج المسكن.

(٣) هل المطبخ حجرة مستقلة أو لا يوجد

(٤) هل الحمام مستقل مع المرحاض لايوجد

(٥) هل المرحاض مستقل مع الحمام لايوجد

(٦) هل المنزل به صرف صحى - نعم لا:

هـ- مادة البناء: (١) طوب أحمر (٢) حجر جيرى (٣) مواد أخرى

(٤) هل الأرضية بلاط أسمنت خشب تراب مواد أخرى

(١) هل المبنى له توصيلة مياه أو بدون توصيلة

و- مياه الشرب: (٢) ما مصدر المياه

ز- الكهرباء: (١) هل المبنى به توصيلة كهرباء أو بدون كهرباءس

تصنيف وجدولة البيانات

عندما يتجمع لدى الباحث كمية كبيرة من الحقائق أو البيانات الإحصائية فى صورة غير منتظمة والتى تم جمعها سواء عن طريق الحصر الشامل Survey كالتعداد السكانى Population Census، أو عن طريق العينات Samples، فإنه يتعذر عليه إستيعاب أو استخلاص النتائج من هذه البيانات قبل تصنيفها وتنظيمها وتلخيصها فى صورة جدولية. ويتوقف تصنيف وجدولة البيانات على طبيعتها والغرض من البحث الذى يسعى الباحث إلى تحقيقه.

ولا توجد طريقة موحدة يمكن الاعتماد عليها فى جدولة البيانات، ولكن هناك قواعد عامة يجب مراعاتها عند تصميم جداول البيانات الإحصائية، وهى:

١ - أن يكون عنوان الجدول محددا لما يتضمنه من بيانات.

٢ - أن ترتب البيانات بالجدول حسب أهيمتها أو تسلسلها الزمنى، مع توضيح وحدات القياس التى جمعت على أساسها.

٣ - أن يوضح المصدر الذى اعتمد عليه فى تكوين بيانات الجدول.

تجميع البيانات:

بعد اكتمال عملية البيانات يتم تجميع المعلومات من الاستمارات الإحصائية ـكشف البحث أو استمارة الاستبيان)، ووضعها غى استمارة خاصة تعرف باستمارة التجميع أو التفريغ Master Sheet، حيث يعطى فيها كل متغير رمزاً

خاصاً به Coding. وعادة ما نحرر استمارة التجميع على ورق المربعات حتى تسهل عملية التفريغ، ويقوم الباحث بتقسيمها إلى محورين: الأول يشمل أنواع المتغيرات التى تندرج تحتها جميع البيانات التى تم جمعها. ويشمل المحور الثانى أرقام أو رموز الاستمارات الإحصائية التى جمع فيها البحث بيانات بحثه. فمثلاً إذا كانت أسئلة الاستمارة الإحصائية تتضمن بيانات عن التركيب العمرى والجنس وحجم الأسرة ودرجة التعليم ممثلة بعدد سنوات الدراسة والمستوى الاقتصادى ممثلاً بالدخل السنوى للمبحوثين، فإن الباحث يقوم بإعطاء كل متغير من هذه المتغيرات رقما خاصا به، كأن يعطى الرقم (١) ليمثل التركيب العمرى، والرقم (٢) ليمثل الجنس والرقم (٣) ليمثل حجم الأسرة ... وهكذا. ويمكن أن ترتب مثل هذه البيانات كما فى الجدول التالى:

جدول رقم (٢-١)

طريقة تجميع المعلومات من الاستمارة الإحصائية

رقم الإستمارة	(١) التركيب العمرى			(٢) الجنس		(٣) حجم الأسرة	(٤) درجة التعلم		(٥) الدخل السنوى
	أقل من ١٥	من ١٥ إلى ٣٠	أكثر من ٣٠	ذكر	أنثى		غير متعلم	متعلم	جنيه
١			٥٣	ذكر		٤		١٦	١٤٤٠
٢		٢٣			أنثى	٢		١٢	٦٦٠
٣			٤٢	ذكر		٣		١٨	١٠٨٠
٤			٦١	ذكر		٥		٩	٨٤٠
٥			٣٦	ذكر		٤		٢٠	١٨٠٠
٦		٢٥		ذكر		٣		٦	٤٢٠

وهكذا فإن الاستمارة رقم (١) تدل على المبحوث الأول الذى يكون عمره ٥٢ سنة وحجم أسرته ٤ أفراد وهو متعلم قضى فى التعليم ١٦ سنة ودخله السنوى

١٤٤٠ جنيهاً. ويطلق على هذا النوع من الاستمارة أسم «جدول التفريغ» حيث يتم تفريغ المعلومات به بعد كتابة عنوانه ووحداته ومصدره.

تصنيف البيانات:

يعتبر تصنيف Classification أو ترتيب البيانات وتقسيمها تقسيماً يسهل إدراك ما بينها من علاقات ويوضح صفاتها ومميزاتها، أساساً لأى أسلوب من أساليب التحليل العلمى ويتم تصنيف المعلومات والبيانات على أسس مختلفة يحددها شكل البحث من جهة والخصائص المميزة للبيانات من جهة أخرى. ويمكن الإشارة هنا إلى خسمة أنواع من التصنيفات ترتب على أساسها البيانات، وهى:

١- التصنيف الأبجدى، ويتم فيه ترتيب مفردات الظاهرة، كترتيب الدول مثلاً، حسب أبجديتها، فمثلاً تسبق دولة أندونيسيا دولة بورما ثم تأتى جامايكا والجزائر .... وهكذا.

٢- التصنيف الزمنى، وفيه ترتب المعلومات حسب الفترات الزمنية (سنوات)، فإذا كان المقصود دراسة تطور قيمة المنتجات الصناعية فى مصر فى الفترة من ١٩٥٧ إلى ١٩٦١، فترتب المعلومات متتالية حسب السنوات. ويوضح ذلك الجدول الآتى:

جدول رقم (٢-٢)

قيمة المنتجات الصناعية فى جمهورية مصر العربية

فى السنوات ١٩٥٧ - ١٩٦١

السنة	قيمة المنتجات (بالآف الجنيهات)
١٩٥٧	٢٩٢٢٤٩
١٩٥٨	٤١٨٥٩٧
١٩٥٩	٤٤٨٧٦٦
١٩٦٠	٤٩٢٧٤٥
١٩٦١	٦٤٧٩٩٢

(المصدر: الإدارة المركزية للإحصاء (مصلحة الإحصاء والتعداد) إحصاء الإنتاج للصناعى)

٣- التصنيف الجغرافى، أى تقسيم العالم إلى وحدات جغرافية كبرى (قارات ومحيطات)، أو تقسيم الدول إلى وحدات جغرافية (أو مقاطعات) والجدول التالى يوضح هذا النوع من التصنيف.

جدول رقم (٢-٣)

مساحات قارات ومحيطات العالم

القارة	المساحة مليون كيلومتر مربع	المحيط	المساحة م.ك٢
أفريقيا	٣٠,٣	الهادى	١٨٣و٤
آسيا	٢٦,٩	الاطلنطى	١٠٦,٧
أوربا والأتحاد السوفيتى	٢٥,٤	الهندى	٧٣,٨
أمريكا الشمالية	٢٤,٣	القطبى الجنوبى	١٩,٧
أمريكا الجنوبية	١٧,٩	القضبى الشمالى	١٢,٤
استراليا ونيوزيلندا	٨,٥		
المجموع	١٣٣,٣	المجموع	٣٨٦,٠

٤- التصنيف الكيفى أو النوعى Qualitative وفيه يتم تقسيم المفردات حسب خصائصها، كأن يصنف الأفراد حسب الحالة الزواجية إلى متزوج، مطلق، أرمل، ولم يتزوج أبداً (أعزب). أو حسب الحالة التعليمية إلى أمى، يقرأ ويكتب ومتعلم. أو تصنف الأقاليم الجغرافية حسب درجات الحرارة إلى أقاليم حارة، أقاليم معتدلة وأقاليم باردة، أو حسب شدة سقوط المطر إلى أقاليم جافة وأخرى رطبة، أو حسب الإرتفاع إلى أقاليم سهلية وأخرى جبلية.

٥- التصنيف الكمى Quantiative، أى إعطاء قيم أو رتب رقمية الخصائص هو وضع الدراسة وترتيبها تصاعدياً أو تنازلياً. ويتم التصنيف على أساس «متغير أحصائس» قابل للقياس الكمى. ثم نوضع بيانات المتغيرات بعد تبويبها فى مجموعات متساوية (أو متقاربة) فى جدول يسمى بالجدول التكرارى. Fre-qucncy Table، كما فى الجدول التالى:

جدول رقم (٢-٤)

توزيع ارتفاعات قمم التلال المحيطة بشمال ويلز

(أخذت هذه البيانات من خريطة أطلس بمقياس ١ : ٢٥,٠٠٠)

فئات الارتفاع (متر)	عدد التلال
أقل من ٢٠٠	٦٢٠
٢٠٠ - ٣٩٩,٩	٦٠٠
٤٠٠ - ٥٩٩,٩	٤٣٠
٦٠٠ - ٧٩٩,٩	٤٠٠
٨٠٠ - ٩٩٩,٩	٤٥٠
١٠٠٠ - ١١٩٩,٩	٢٥٠
١٢٠٠ - ١٣٩٩,٩	٢٥٠
١٤٠٠ - ١٥٩٩,٩	٢٠٠
أكثر من ١٦٠٠	٦٠١

وتجدر الإشارة إلى أنه قد يشتمل التصنيف على أكثر من نوع واحد من الأنواع السابقة. فقد يكون التصنيف زمنياً وجغرافياً فى نفس الوقت، كأن يتضمن الجدول بيانات ظاهرة ما لعدد من السنوات موزعة حسب الوحدات الجغرافية للمنظمة موضع الدراسة. وقد يكون التصنيف نوعياً وجغرافياً فى نفس الوقت أيضاً، كأن يتضمن الجدول بيانات خصائص الظاهرة أو أقسامها الفرعية وذلك بالنسبة لكل وحدة جغرافية للمنظمة. وقد يكون التصنيف نوعياً وكمياً فى ذات الوقت إذا تضمن الجدول بيانات عن حجم أو كمية الظاهرة وذلك بالنسبة لكل قسم أو صفة من صفات نفس الظاهرة.

الجدولة اليدوية للبيانات الإحصائية:

يقصد بجدولة البيانات الإحصائية أو ما يعرف «بتبويب البيانات» وضعها أو

عرضها فى أشكال جدولية مختلفة لغرض التبسيط وتلخيص خصائص المعلومات، وبطريقة تسهل دراستها وتحليلها باستخدام المقاييس الإحصائية والأساليب الكمية المختلفة للوصول إلى نتائج واتخاذ القرارات التى يمكن تعميمها على المجتمع الإحصائى. وتختلف طرق العرض الجدولى اليدوية (غير الآلية) باختلاف الأسلوب المستخدم، كما تتنوع الجداول الإحصائية باختلاف طبيعة البيانات المراد جدولتها. ويمكن تقسيم الجداول الإحصائية إلى نوعين رئيسيين هما: الجداول العادية، والجداول التكرارية.

الجداول العادية:

تقسم الجداول العادية عموماً إلى ثلاثة أنواع:

١- الجداول البسيطة، وهى التى تتمثل فيها الظاهرة وخصائصها الكمية الممثلة بالقيم والأرقام، كما هو فى الجدول التالى:

جدول رقم (٢-٥)

توزيع ٢٠ أسرة حسب الحجم

حجم الأسرة «فرد»	عدد الأسرة
٢	٢
٣	٢
٤	٦
٥	٥
٦	٢
٧	٢
٨	١
الجملة	٢٠

٢- الجداول المركبة، وهى التى تتمثل فيها الظاهرة وخصائصها الممثلة بالأرقام داخل عدة أعمدة الجدول كأن نحدد عدد سكان الحضر والريف فى الوحدات الجغرافية الكبرى لجمهورية مصر العربية.

جدول رقم (٢-٦)
التوزيع الجغرافى لسكان كل من الريف
والحضر فى مصر (تعداد ١٩٦٠)

المحافظات	عدد السكان بالآلاف		المجموع
	حضر	ريف	
محافظات الحضر	٥٥٨٢	-	٥٨٨٢
الوجه البحرى	٢٠٥٠	٨٧٨٨	١٠٨٣٨
الوجه القبلى	١٨٩٨	٧٣٣١	٩٢٢٩

٣- الجداول المزدوجة، وهو ذلك النوع من الجداول الذى يبين توزيع البيانات حسب صفتين فى نفس الوقت. وفيه تمثل الأعمدة تقسيمات أحد الصفتين، بينما تمثل الصفوف تقسيمات الصفة الأخرى. ومن أوضح الأمثلة لهذا النوع جداول Input - Output Taples مثل الجداول الخاصة بالميزانيات المصرفية (الواردات - المصروفات) أو جداول الإنتاج والإستهلاك، أو جداول توزيع الدخل على فئات عمر السكان. كما يظهر فى الجدول التالى:

جدول رقم (٢-٧)
فئة الدخل (بالجنيهات) والأعمار بالسنين

الدخل / السن	٣-	٥-	٧-	٩-	١١-	المجموع
٢٠-	٤	٢				٦
٣٠-		٤	٧			١١
٤٠-			٥	٣		٨
٥٠-			٠	٣	٢	٥
المجموع	٤	٦	١٢	٦	٢	٣٠

الجداول التكرارية: Frequency Tables

عند تلخيص كمية كبيرة من البيانات ذات الصفة الكمية المتغيرة فإنه من المفيد ترتيبها أو توزيعها على فئات وتحديد عدد المفردات التى تنتمى لكل فئة. والجدول المنظم على صورة فئات يقابل كل فئة تكرار معين يسمى جدول التوزيع التكاراى كما فى الجداول التالى:

جدول رقم (٢-٨)

توزيع تكرارى لأوزان ١٠٠ طالب

الأوزان (كيلو جرام)	عدد الطلبة (التكرار)
٦٠ - ٦٢	٥
٦٣ - ٦٥	١٨
٦٦ - ٦٨	٤٢
٦٩ - ٧١	٢٧
٧٢ - ٧٤	٨
المجموع	١٠٠

وتتلخص القواعد العامة لتكوين جدول التوزيع التكرارى فى الخطوات التالية:

١- تحديد أكبر وأقل قيمة فى البيانات ومنها توجد المدى (الفرق بين أكبر رقم وأقل رقم).

٢- تقسيم المدى إلى عدد مناسب من الفئات المتساوية الطول. وإذا لم يكن ذلك ممكناً تستخدم فئات ذات أطوال مختلفة أو فئات مفتوحة. ويؤخذ عدد الفئات عادة بين ٥، ٢٠ فئة حسب البيانات المتاحة. وتختار الفئات أيضاً بحيث يتفق مركز الفئة مع المشاهدات الفعلية. وهذا يؤدى إلى التقليل من أخطاء التجميع عند إجراء مزيد من المعالجة الإحصائية.

٣- تحديد عدد المشاهدات (أو التكرارات) التي تقع في كل فترة فئة وأحسن طريقة لأداء ذلك هو إستخدام كشف الحزم أو العلامات.

وتسمى البيانات المنمطة والملخصة في جداول التوزيع التكراري بالبيانات المجمعة. وعلى الرغم من أن عملية التجميع تؤدى بشكل عام إلى ضياع كثير من التفصيلات الأصلية للبيانات فان الفائدة الهامة منها هى الصورة العامة التى يمكن الحصول عليها والتى تسهل تحليل وإستخلاص النتائج والحقائق من البيانات.

اختيار الفئات:

لا توجد قاعدة ثابتة لتحديد أطوال وعدد الفئات، حيث أن ذلك يتوقف على طبيعة البيانات وعلى كمية الاختلاف بين المفردات وعلى الدقة المطلوبة فى العمليات الحسابية. ولكن عند تحديد فترة (طول) الفئة فانه يحسن الأخذ فى الاعتبار العدد الكلى للقيم. فكبر فترة الفئة يؤدى إلى الحصول على عدد من الفئات أقل مما لو كانت فترة الفئة أصغر. كما أنه عند تحديد عدد الفئات ينبغى اختيار عدد قليل من الفئات يؤدى إلى تسهيل العمليات الحسابية. ولتحديد عدد الفئات تستخدم إحدى الطرق الآتية:

١- معادلة Yule = عدد الفئات = ٢,٥ $\sqrt[4]{ن}$

حيث ن تمثل عدد مفردات البيانات. وتصلح هذه المعادلة عندما يكون حجم العينة أقل من ١٠٠٠ مفردة.

٢- معادلة استرجس Sturges : عدد الفئات = ١+ (٣,٣٢٢ + لو ن) حيث لو ن تمثل لوغاريتم عدد القيم. وحيث أن هذه الصيغة تعطى عدد به كسور عشرية، فيجوز هنا تقريب الناتج للحصول على مقدار صحيح لعدد الفئات. وتصلح هذه المعادلة عندما تزيد مفردات البيانات عن ١٠٠٠ مفردة.

وتلعب الخبرة فى مجال تحديد عدد الفئات دوراً كبيراً. فلا يجب تقليل عدد الفئات حتى لا يؤدى ذلك إلى فقد بعض التفاصيل الموجودة فى البيانات الخام، كذلك لا يجب زيادة عدد الفئات بحيث يصعب معها عملية الدراسة والتحليل

بعد ذلك. على أنه تجدر الإشارة إلى أن أنسب الجداول التكرارية هو الذى يحتوى على عدد من الفئات يتراوح بين ٨ إلى ١٢ فئة. بل يجب أن لا يتضمن الجدول على أقل من ٦ فئات ولايزيد عن ٢٠ فئة.

فترة وحدود الفئات: Class (Cell) Interval and Limits

يعرف الرمز الذى يعبر عن الفئة ٦٠-٦٢ فى الجدول رقم (٢-٨) بفترة الفئة، كما يطلق على الرقمين ٦٠، ٦٢ حدود الفئة. ويسمى الرقم الأصغر (٦٠) بالحد الأدنى للفئة، والرقم الأكبر (٦٢) بالحد الأعلى للفئة. وفى معظم التوزيعات التكرارية تكون فترات الفئات متساوية. ويمكن تحديد فترة الفئة بالصورة الآتية:

$$\text{فترة (طول) الفئة} = \frac{\text{(الفرق بين أكبر وأصغر قيمة فى البيانات}}{\text{عدد الفئات}}$$

وحيث أن هذه الصيغة تعطى طول فترة به كسور عشرية، فمن المستحسن تقريب الناتج للحصول على مقدار (أو رقم) صحيح لفترة الفئة.

ويمكن أيضاً تقدير فترة الفئة مباشرة بدون حساب عدد الفئات وذلك عن طريق العلاقة بين فترة الفئة والإنحراف المعيارى للبيانات. فمن المعروف من هذه العلاقة أن فترة الفئة والإنحراف المعيارى للبيانات. فمن المعروف من هذه العلاقة أن فترة الفئة عادة تتراوح عن ١/٤ و ١/٢ الإنحراف المعيارى للقيم، أو بمعنى آخر لاتقل فترة الفئة عن ١/٤ عن الإنحراف المعيارى ولا تزيد عن ١/٢ الإنحراف المعيارى. ولقد تم حساب الإنحراف المعيارى للبيانات المستمرة بمعرفة كل من عدد المفردات فى العينة والمدى بينها، ووضع فى جداول خاصة يطلق عليها اسم جداول Tippet كما فى الجدول التالى:

جدول رقم (٢-٩)

قيمة الإنحراف المعيارى للأحجام المختلفة من العينات

حجم العينة	٥٠	٧٥	١٠٠	١٥٠
الإنحراف المعيارى	٠,٢٢٢ى	٠,٢٠٨ى	٠,١٩٩ى	٠,١٨٩ى
حجم العينة	٢٠٠	٣٠٠	٥٠٠	٧٠٠
الإنحراف المعيارى	٠,١٨٢ى	٠,١٧٢ى	٠,١٦٤ى	٠,١٥٩ى

ى = المدى

فعلى سبيل المثال إذا كان العينة ١٠٠ والمدى بين المفردات ٤٦ فإن فترة الفئة يمكن تحديدها كمايلى:

الإنحراف المعيارى للبيانات = ٠,١٩٩ × ٤٦ = ٩,١٥٤

∴ فترة الفئة تترواح بين (٩,١٥٤) ، (٩,١٥٤) أى ما بين ٢ ، ٥ تقريباً.

وبعد اختيار عدد الفئات وتحديد فتراتها يلى ذلك تحديد بداية ونهاية كل فئة وهو ما يسمى بحدود الفئات Class Boundaries. وتوضح حدود الفئات بحيث يكون الحد الأدني للفئة الأولى أقل قليلاً من أقل قيمة فى البيانات، أما الحد الأعلى للفئة فيجب أن ينتهى قبل بداية الفئة التالية وهكذا بالنسبة لبقية الفئات حتى لا تتداخل حدود الفئات ويصبح من الصعب توزيع البيانات.

ومن الناحية النظرية تتضمن فترة الفئة كل القياسات بين حدى الفئة إذا كان المتغير موضع الدراسة من نوع المتغيرات غير المستمرة (الوثابة) التى لا تأخذ فيها كسرية كما فى الجدول رقم (٢-٨). فيما أن الأوزان سجلت لأقرب كيلو جرام فإن فترة الفئة ٦٠ - ٦٢ تشتمل على كل القياسات من ٥٩,٥ إلى ٦٢,٥، وهذه الأرقام تسمى بالحدود الحقيقية للفئة. والرقم الأصغر ٥٩,٥ هو الحد الأدنى الحقيقى للفئة التالية والقسمة على ٢. ويجب أن تكتب الفئات بحيث لا تتطابق

مع أحد القيم الفعلية. فإذا كانت الفئات المختلفة بالعمود الأول فى الجدول رقم (٢-٨) مكتوبة بالصور ٦٠-٦٢، ٦٢-٦٥، ٦٥-٦٨، وكان لدينا القيمة ٦٥ فإنه يكون من الصعب تقرير ما إذا كانت تنتمى إلى الفئة ٦٢-٦٥ أو الفئة ٦٥ - ٦٨.

وفى حالة إذا كان المتغير قيد البحث من المتغيرات المستمرة (المتصلة) فيكون من الخطأ كتابة حدود الفئات بالصورة التى تكتب بها فى حالة فئات المتغير غير المستمر. ويمكن كتابة حدود الفئة السابقة لها دون أن يحدث تداخل أو تترك ثغرات بين الفئات. وتكتب الفئات كمايلى: ٦٠ إلى أقل من ٦٣، ٦٢ إلى أقل ٦٥، ٦٥ من ٦٨ .... وهكذا. وللاختصار تكتب الحدود الدنيا للفئات وتترك حدودها العليا، إلا أنه فى هذه الحالة يجب تحديد نهاية الفئة الأخيرة كمايلى: ٦٠ -، ٦٢-، ٦٥-، ...، ٧١ إلى أقل من ٧٤.

مركز الفئة : Class mid-point

مركز الفئة هو منتصف فترة الفئة ونحصل عليه بجمع الحد الأدنى والحد الأعلى للفئة وقسمة المجموع على أثنين. فمركز الفئة ٦٠-٦٢ هو $\frac{٦٠ + ٦٢}{٢}$.

ويسمى مركز الفئة أيضاً بمنتصف الفئة. وعند التحليل الاحصائى للبيانات الموزعة تكراريا داخل فئات فإنه يفترض أن جميع التكرارات الموجودة داخل فترة فئة تأخذ قيما تتطابق مع مركز الفئة. وعلى ذلك فان جميع الأوزان داخل الفئة ٦٠-٦٢ فى الجدول رقم (٢-٨) تعتبر كما لو أنها ٦١ كيلو جرام.

تكوين الجدول التكرارى:

بعد اختبار عدد الفئات وتحديد فتراتها وحدودها نبدأ فى تكوين جدول يشتمل على عدد من الصفوف مساو لعدد الفئات، ثم توزع البيانات الخام على الفئات المختلفة حسب قيمتها وحصر عدد المفردات الموجودة داخل فترة كل فئة باحدى الطرق الآتية:

١ - طريقة المنظومة Array، وفيها يتم ترتيب وتنظيم البيانات إما تنازليا أو تصاعديا

حسب قيمتها مما يسهل تحديد عدد المفردات داخل كل فئة. وعلى الرغم من ذلك فان هذه الطريقة لا يمكن تطبيقها على البيانات ذات المفردات الكثيرة إذ تصبح غير عملية فى هذا الشأن.

٢- طريقة الحزم أو العلامات Tally marks، وفيها يتم توزيع مفردات البيانات مباشرة على الفئات وذلك بوضع (/) أمام الفئة المناسبة ثم نحصر عدد هذه العلامات عند نهاية تفريغ جميع مفردات البيانات. ويلاحظ أثناء التوزيع أننا نضع كل أربع علامات بجوار بعضها أما العلامة الخامسة فتوضع فوق الأربع علامات السابقة وبذلك تكون حزمة (////) (أنظر الجدول رقم ٢-١١). والغرض من ذلك هو تسهيل عد العلامات داخل كل فئة. ويطلق على عدد العلامات أمام كل فئة إسم «تكرار الفئة Frequency» ويرمز له بالرمز (ك). وعند العرض النهائى للتوزيع التكرارى يحذف عادة كشف الحزم (أو العلامات) من جدول ترتيب البيانات.

وسوف نوضح خطوات عمل جدول التوزيع التكرارى للبيانات التالية والتى تمثل أطوال مائة رافد نهرى (بالمتر) لأحد أحواض التصريف النهرى.

جدول رقم (٢-١٠)

أطوال مائة رافد نهرى (بالأمتار)

٨٥٠	١٠٨٠	٨٢٠	٨٥٠	١١١٠	٩٩٠	١١٦٠	١٠٠٠	١١٢٠	٧٥٠
١٠٢٠	٩٢٠	١٠٦٠	٩٨٠	١٣٣٠	١٠٢٠	١٠٣٠	١١٨٠	٩١٠	٩٤٠
١٠٧٠	١٠٠٠	١٢١٠	٩٤٠	٧٧٠	١٠٨٠	٥٧٠	١٠٠٠	١٠٩٠	١١٥٠
٨٨٠	١٠١٠	٩٨٠	١٠٢٠	٩٧٠	٨٧٠	٨٨٠	١١١٠	٦٧٠	١٠٤٠
١٠٣٠	١٠١٠	٩١٠	١٢٠٠	١٠٦٠	٨٠٠	١٠٧٠	٨٥٠	٩٣٠	٩٠٠
١٠٦٠	٨٩٠	٩٨٠	٨٣	٩٨٠	١١٢٠	١٠٧٠	١٢٧٠	١٠٠٠	١٠٩٠
٨٧٠	١٠٢٠	٩٠٠	١٠٠٠	٩٣٠	١١٩٠	٩٤٠	٨٥٠	١١٧٠	٧٩٠
١٠٤٠	١٢٢٠	١٠٥٠	٩٨٠	٧٢٠	٩٣٠	٩٤٠	١٤٢٠	١٠٩٠	٩٥٠
١٠٧٠	١٠٣٠	١٠٩٠	١٠٠٠	٩٧٠	١٠٧٠	١٠٤٠	١٠٢٠	٧٩٠	١٠٤٠
٨٨٠	٩٨٠	٧٦٠	١٠٢٠	١١٠٠	١٠٢٠	١٠٧٠	٨٣٠	٩٦٠	١٠٦٠

أولاً: تحديد المدى الذى تتغير فيه الأطوال، أى تحديد أكبر وأقل قيمة فى البيانات وإيجاد الفرق بينهما. وبالنسبة للبيانات السابقة فإن أكبر طول هو ١٤٢٠ وأصغر طول هو ٥٧٠ وبهذا يكون المدى = ١٤٢٠ – ٥٧٠ = ٨٥٠

ثانياً: تحديد عدد الفئات. حيث أن عدد القيم أقل من ١٠٠٠ فإننا نستخدم معادلة Yule لإيجاد عدد الفئات كما يلى:

عدد الفئات = ٢,٥ $\sqrt[٤]{١٠٠}$ = ٩و٧ فئة

ثالثاً: تحديد فترة الفئة، إذا استخدمنا فئات فإن

فترة الفئة = $\frac{\text{المدى}}{\text{عدد الفئات}}$ – $\frac{٨٥٠}{٨}$ = ١٠٦,٢٥ متراً

وإذا استخدمنا ٩ فئات فإن

فترة الفئة = $\frac{\text{المدى}}{\text{عدد الفئات}}$ = $\frac{٨٥٠}{٩}$ = ٩٤,٤٤ متراً

ولتجنب وجود كسور عشرية فى مراكز الفئات فإنه من الأفضل أن تكون فترة الفئة رقما صحيحاً وليكن ١٠٠ متراً. وكذلك فإنه من الملائم اختيار مراكز الفئات عند ٥٩٥، ٦٩٥، ٧٩٥ ...... ولهذا فإن الفئات من الممكن أن تكون ٥٥٠ – ٦٤٠، ٦٥٠ – ٧٤٠.. والجدول التالى يوضح عملية تكوين الجدول التكرارى.

جدول رقم (٢-١١)

جدول التوزيع التكرارى لأطول ١٠٠ رافد نهرى

فترة الفئة (الطول بالمتر)	حدود الفئة	مركز الفئة	الحزم	التكرار (ك)
٥٥٠-٦٤٠	٥٤٥-٦٤٥	٥٩٥	/	١
٦٥٠-٧٤٠	٦٤٥-٧٤٥	٦٩٥	//	٢
٧٥٠-٨٤٠	٧٤٥-٨٤٥	٧٩٥	𝍸 𝍸	٩
٨٥٠-٩٤٠	٨٤٥-٩٤٥	٨٩٥	𝍸 𝍸 𝍸 𝍸 //	٢٢
٩٥٠-١٠٤٠	٩٤٥-١٠٤٥	٩٩٥	𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 //	٣٣
١٠٥٠-١١٤٠	١٠٤٥-١١٤٥	١٠٩٥	𝍸 𝍸 𝍸 𝍸 //	٢٢
١١٥٠-١٢٤٠	١١٤٥-١٢٤٥	١١٩٥	𝍸 ///	٨
١٢٥٠-١٣٤٠	١٢٤٥-١٣٤٥	١٢٩٥	//	٢
١٣٥٠-١٤٤٠	١٣٤٥-١٤٤٥	١٣٩٥	/	١
	المجموع الكلى للمفردات = مجموع التكرارات			١٠٠

ويلاحظ أننا استـخـدمنا فى كل من الجـدولين رقم (٢-٧) ورقم (٢-٨) فـئـات ذات فـتـرات مـتـسـاوية. ويطلق على التـوزيع التكرارى من هذا النـوع اسم «التوزيع المنتظم». ويفضل فى كل الحالات استخدام فئات متساوية وذلك تسهيلاً للعمليات الحسابية. وهناك نوع آخر من التوزيعات التكرارية يسمي بالتوزيع غير المنتظم، وفيه تكون أطوال الفئات غير متساوية وذلك بسبب تركز عدد كبير من مفردات البيانات فى مدى ضيق نسبياً وانتشاراً العدد الأصغر منها على مدى واسع. ومن أمثلة هذا النوع من التوزيعات توزيع ملكية الأرض حسب فئات المساحة.

وتجدر الإشارة إلى أنه يوجد بصفة عامة نوعان من جداول التوزيع التكرارى حسب نوع التوزيع هما: مايعرف «بالجدول المفتوح» وفيه يكون الحد الأدنى للفئة والحد الأعلى للفئة الأخيرة غير محدد. وهو فى هذه الحالة يكون مفـتـوح من طرفيه. ويكون الجدول مفتوحاً من طرفه الأعلى إذا كانت بداية الفئة الأولى غير

محددة. ويكون مفتوحاً من طرفه الأدنى إذا كانت نهاية الفئة الأخيرة غير محددة ويكون مفتوحاً من طرفه الأدنى إذا كانت نهاية الفئة غير محددة. أما النوع الثانى من الجداول فيعرف باسم «الجدول المقفول» وفيه يكون الحد الأدنى للفئة الأولى والحد الأعلى للفئة الأخيرة محددين كما في الجدولين رقم (٢-٨) و (٢-١١).

التوزيع التكرارى النسبى Relative Frequency Distribution

يعرف التكرار النسبى لفئة بأنه عبارة عن تكرار الفئة مقسوماً على التكرار الكلى لجميع الفئات، وعادة يعبر عنه كنسبة مئوية. ويستخدم التكرار النسبى فى حالة مقارنة توزيعين تكراريين يختلفان فى عدد مفردات كل منهما إذ أن مقارنة التكرارات المطلقة لكل فترة فى التوزيعين تكون غير صحيحة بسبب اختلاف حجم عينة كل منهما. فعلى سبيل المثال، فإن التكرار النسبى للفئة ٥٥٠ – ٦٤٠ فى الجدول رقم (٢-١١) هو ١/١٠٠ = ١٪ وإذا استبدلنا التكرارات في الجدول التكرارى (جدول رقم ٢-١١) بما يقابلها من التكرارات النسبية فإن الجدول الناتج يسمى بالتوزيع التكرارى النسبى أو جدول التكرارات النسبية كما يبدو من الجدول التالى:

جدول رقم (٢-١٢)

التوزيع التكرارى النسبى لأطوال ١٠٠ رافد نهرى

مركز الفئة	التكرار (ك)	التكرار النسبى
٥٩٥	١	١ + ١٠٠ = ٠,٠١ (١٪)
٦٩٥	٢	٢ + ١٠٠ = ٠,٠٢ (٢٪)
٧٩٥	٩	٩ + ١٠٠ = ٠,٠٩ (٩٪)
٨٩٥	٢٢	٣٢ + ١٠٠ = ٠,٢٢ (٢٢٪)
٩٩٥	٣٣	٣٢ + ١٠٠ = ٠,٣٢ (٣٢٪)
١٠٩٥	٢٢	٢٢ + ١٠٠ = ٠,٢٢ (٢٢٪)
١١٩٥	٨	٨ + ١٠٠ = ٠,٠٨ (٨٪)
١٢٩٥	٢	٢ + ١٠٠ = ٠,٠٢ (٢٪)
١٣٩٥	١	١ + ١٠٠ = ٠,٠١ (١٪)

التوزيع التكرارى المتجمع Cumulative Frequency Distribution

يسمى مجموع التكرارات المقابلة لجميع القيم الأقل من الحد الأعلى لفئة معينة بالتكرار المتجمع إلى هذه الفئة والمتضمن تكرارها أيضاً. وعلى سبيل المثال ففى الجدول رقم (٢-١١) فإن التكرار المتجمع فى الفئة ٨٥٠ - ٩٤٠ والمتضمن تكرارها أيضاً هو ١+ ٢+ ٩+ ٢٢ = ٣٤. وهذا يعنى أن ٣٤ رافداً نهرياً تقل أطوالاً عن ٩٤٠ متراً. والجدول الذى يمثل التكرارات المتجمعة يسمى بجدول التوزيع المتجمع أو جدول التكرارات المتجمعة كما هى الحال فى الجدول رقم (٢-١٣).

وفى بعض الأحيان قد يكون من المرغوب فيه الحصول على التوزيع التكرارى المتجمع لجميع القيم الأكبر أو المساوية للحد الأدنى لكل فئة. ويسمى التوزيع فى هذه الحالة بالتوزيع المتجمع على أساس «أكبر من» أو التكرار المتجمع.

جدول رقم (٢-١٣)

التوزيع التكرارى المتجمع الصاعد والهابط لأطوال ١٠٠ رافد نهرى

التوزيع التكرارى		التوزيع التكرارى المتجمع الصاعد		التوزيع التكرارى المتجمع الهابط	
الفئة	التكرار	حدود الفئة	التكرار المتجمع	حدود الفئة	التكرار المتجمع
٥٥٠-٦٤٠	١	أقل من ٦٤٠	١	أكبر من ٥٥٠	١٠٠
٦٥٠-٧٤٠	٢	أقل من ٧٤٠	١ + ٢ = ٢	أكبر من ٦٥٠	١٠٠ - ١ = ٩٩
٧٥٠-٨٤٠	٩	أقل من ٨٤٠	٣ + ٩ = ١٢	أكبر من ٧٥٠	٩٩ - ٢ = ٩٧
٨٥٠-٩٤٠	٢٢	أقل من ٩٤٠	١٢ + ٢٢ = ٣٤	أكبر من ٨٥٠	٩٧ - ٩ - ٨٨
٩٥٠-١٠٤٠	٣٢	أقل من ١٠٤٠	٣٤ + ٣٣ = ٦٧	أكبر من ٩٥٠	٨٨ - ٢٢ = ٦٦
١٠٥٠-١١٤٠	٢٢	أقل من ١١٤٠	٦٧ + ٢٢ = ٨٩	أكبر من ١٠٥٠	٦٦ - ٣٣ = ٣٣
١١٥٠-١٢٤٠	٨	أقل من ١٢٤٠	٨٩ + ٨ = ٩٧	أكبر من ١١٥٠	٣٣ - ٢٢ = ١١
١٢٥٠-١٣٤٠	٢	أقل من ١٣٤٠	٩٧ + ٢ = ٩٩	أكبر من ١٢٥٠	١١ - ٨ = ٣
١٣٥٠-١٤٤٠	١	أقل من ١٤٤٠	٩٩ + ١ = ١٠٠	أكبر من ١٣٥٠	٣ - ٢ = ١

الهابط (النازل). بينما يسمى التوزيع المتجمع على أساس «أقل من» بالتكرار المتجمع الصاعد وهو يمثل مجموع القيم التى تقل فى قيمتها عن الحد الأعلى للفئة، أى تساوى تكرار الفئات السابقة لفئة ما مضافاً إليها تكرار الفئة نفسها.

وكما ذكرنا سابقاً يمكن استبدال التكرار المتجمع المطلق فى الجدول السابق إلى تكرار متجمع نسبى (أو تكرار متجمع مئوى) Relative Cumulative Frequency وذلك بقسمة التكرار المتجمع فى كل فئة على التكرار الكلى. فمثلاً. التكرار المتجمع النسبى للأطوال أقل من ١٠٤٠ متراً. هو $\frac{٦٧}{١٠٠}$ = ٦٧٪

وهذا يعنى أن ٦٧٪ من الروافد أطوالها أقل من ١٠٤٠ متراً. وإذا وضعت التكرارات المتجمعة النسبية فى جدول فإنها تسمى بالتوزيع التكرارى المتجمع النسبى أو بالتوزيع المتجمع للنسب المئوية.

العرض البيانى للبيانات الإحصائية

يعتمد أسلوب العرض البيانى على ترجمة المعلومات وتلخيص البيانات الإحصائية (المبوبة وغير المبوبة) ووضعها فى صورة أشكال بيانية أو فى هيئة رسوم تصويرية تسهل فهم واستيعاب الخصائص والاتجاهات والعلاقات المختلفة والمتشابكة للظواهر الجغرافية موضع الدراسة. وتبعاً لذلك فان الأشكال والرسوم البيانية تعد خير وسيلة للتعبير وتوصيل المعلومات، كما يمكن أن نعتبرها لغة ثانية يشرح بها الباحث موضوع بحثه دون أن يجهد القارئ أو المشاهد فى استخلاص الحقائق من الجداول والأرقام. وتمتاز الأشكال والرسوم البيانية والتصويرية بأنها تعطى فكرة سريعة للناظر إليها من أول وهلة، بينما لا يظهر هذا الأثر إذا مانظرنا إلى بيانات رقمية فى جدول أو إحصائية. لكل ذلك فإننا - بحق - يمكن أن نقول أن العرض البيانى هو روح البيانات وسبيل إلى الوصول إلى ما تخبؤه من معلومات.

العرض البيانى للبيانات الخام (غير المبوبة)

تختلف وتتعدد طرق وأساليب العرض البيانى للبيانات الخام (غير المبوبة)،ولكنها تنحصر فى ثلاث مجموعات رئيسية هى: مجموعة الطرق البيانية لتمثيل التغير فى مكونات عناصر الظاهرة والمجموع الكلى لها والطرق البيانية بالرسوم التصويرية. وفيما يلى دراسة تفصيلية لكل مجموعة على حدة.

أولاً: الطرق البيانية لتمثيل العلاقات بين كميات المتغيرات:

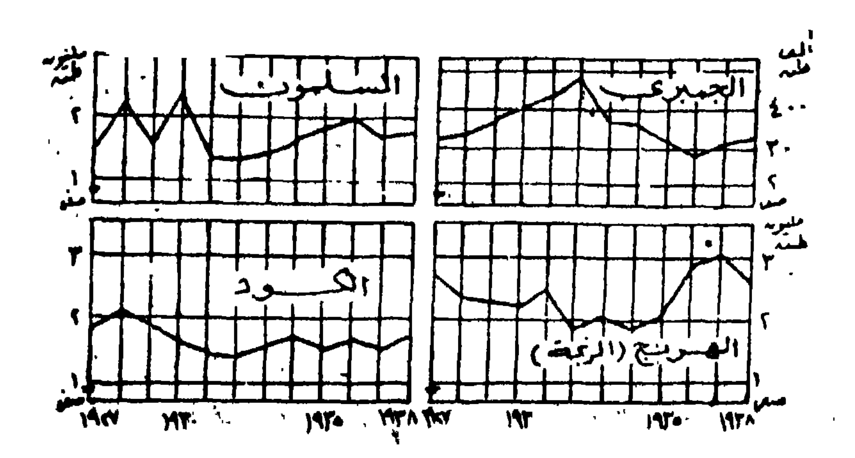
يمكن تقسيم هذه الطرق، حسب البيانات الإحصائية المتاحة للمتغيرات موضع

الدراسة إلى قسمين: القسم الأول يختص ببيان العلاقات بين الكميات المختلفة لعدة متغيرات، والقسم الآخر يهتم بإظهار تطور مقدار (قيم) المجموع الكلى لمتغير واحد أو عدة متغيرات. ويتضمن القسمان مجموعة غير قليلة من أساليب التمثيل البيانى سنتعرض لشرح أكثرها أهمية فى تحقيق الغرضين السابقين، مثل الخطوط البيانية الحسابية واللوغاريتمية، أشكال الإنتشار، الأعمدة البيانية البسيطة، والرسوم البيانية الحجمية والرسوم الدائرية، فيمايلى:

١ - الخط البيانى البسيط Simple-Line graph

تستخدم الخطوط البيانية البسيطة فى تمثيل التغير من فترة إلى أخرى للظاهرة الواحدة أو لبيان علاقة متغيرين وغالبا ما يكون أحد هذين المتغيرين هو الزمن الذى يعتبر متغيراً مستقلاً. ويبين التغير أو العلاقة بمنحنى يظهر منه ضعف أو شدة التغير من فترة إلى أخرى ويكون ذلك ما يعرف بالسلسلة الزمنية، أو يوضح اتجاه العلاقة بين متغيرين أو أكثر.

ولقد جرت العادة عند التمثيل البيانى للسلاسل الزمنية التى توضح مقادير أو كميات المتغيرات لعدد من السنوات أن يكون المحور الأفقى «السنى» ممثلاً للمتغير المستقل «الزمنى» والمحور الرأسى «الصادى» للمتغير التابع الظاهرة موضع الدراسة. وفى ضوء البيانات المتاحة يختار مقياس رسم ملائم لأبعاد المسطح المخصص لعملية التمثيل البيانى حتى يمكن توقيع كل قيم المتغير التابع على الرسم. فيقسم المحور الرأسى إلى وحدات حسابية بادئين بالصفر ومنتهين بقيمة أكبر من أكبر قيمة تمثل المتغير التابع. ويختار كذلك مقياس مناسب للمتغير المستقل «الزمن» على المحور الأفقى ويجب مراعاة عدم وجود تفاوت كبير فى الأبعاد القياسية للمتغيرين حتى لايؤدى ذلك إلى نقص الدقة فى تمثيل بيانات المتغيرين. ويتم رسم المنحنى من خلال توقيع جميع القيم على الرسم فى شكل نقط تحدد كل منها باحداثين (أى على حسب كل نقطة عن المحورين) ثم توصل مواقع القيم فتعطى لنا الشكل المطلوب (الخط البيانى) كما فى الشكل رقم (٣-١).



شكل رقم (٣-١)

تطور إنتاج الثروة السمكية فى كندا فى الفترة من ٢٧ - ١٩٣٨

(طريقة الخطوط البيانية البسيطة)

ويجب ملاحظة أنه عند رسم الخطوط البيانية لظاهرة متغيرة بانتظام أو تدريجيا أن يكون الخط البيانى منحنياً، مثل الخط البيانى الذى يوضح المتوسط الشهرى لدرجة الحرارة خلال شهور السنة فى مدينة الاسكندرية. أما إذا كانت الظاهرة لا تتغير بانتظام مثل كمية الأمطار أو كمية إنتاج أحد المحاصيل الزراعية فإنه يجب أن يكون التوصيل بين النقط على شكل خط منكسر. وفى بعض الحالات قد لا يبدأ المقياس على المحور الرأسى بالصفر ولكن يبدأ بقيمة أكبر تبعا لأن البيانات المراد تمثيلها بيانياً بخط بيانى تبدأ بقيمة بعيدة عن الصفر ولكن تقترب من بعضها بمدى صغير، فإذا ما أخذنا وحدات قياسية تناسب أصغر وأكبر رقم بادئين بالصفر فإن ذلك سيؤدى إلى وجود فراغ كبير غير مستخدم يقع بين الصفر وأصغر رقم موجود مما يترتب عليه أن يجعل الخط البيانى محصوراً فى أعلى جزء من الرسم وهذا شئ غير مستحب أو مرغوب فيه. وللتغلب على ذلك فإننا نحاول أن نضغط المسافة على المحور الرأسى بين الصفر وأصغر قيمة بأن نكسر المحور الرأسى بخطين مائلين بعد نقطة الصفر على المحور الرأسى.

وفى بعض الأحيان إذا كان المطلوب تمثيل بيانات متغير واحد فى أكثر من مكان لابراز العلاقات والاختلافات المكانية لهذا المتغير، أو تمثيل بيانات متغيرين أو أكثر فى مكان واحد لابراز خصائص المتغيرات موضع البحث لهذا المكان، فإنه يمكن رسم وتوقيع هذه البيانات بأكثر من خط بيانى بسيط فى شكل بيانى واحد يعرف باسم « الخطوط البيانية المتعددة أو البوليجراف Polygraphs» .والقاعدة الأساسية لإنشاء البوليجراف تتلخص فى رسم محورين رأسيين – محور أيمن وآخر أيسر – كل منهما يقسم إلى تقسيمات تختص ببيانات أحد المتغيرين.

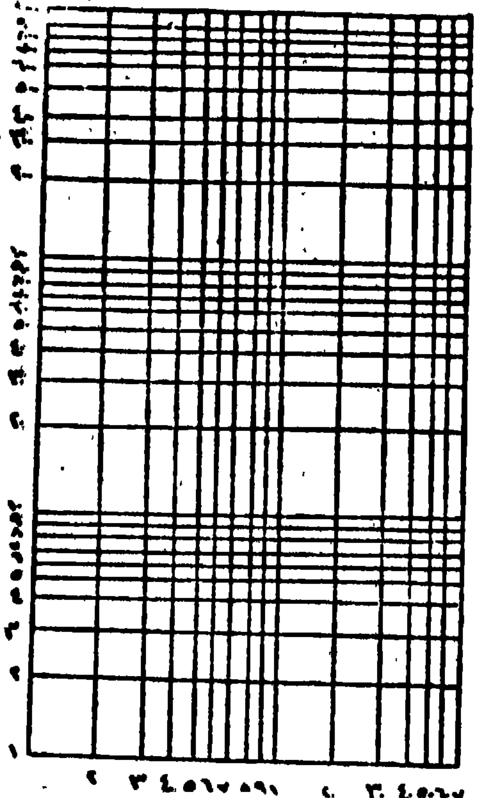
ويستخدم هذا النوع من الرسوم البيانية على نطاق واسع لتوضيح العلاقات بين بيانات العناصر المناخية، مثل بيان العلاقة بين درجة الحرارة وكمية التبخر أو بين كمية الأمطار وعدد الأيام الماطرة، أو بين بيانات أحد العناصر المناخية وبيانات ظاهرات جغرافية طبيعية أخرى مثل التضاريس والنبات، أو ظاهرات بشرية مثل عدد المصطافين أو عدد السياح فى دولة ما، أو كمية استهلاك المياه فى مدينة ما، أو حتى نوع العمل الذى يقوم به الأفراد فى مجتمع ما. وبصفة عامة فإن هناك علاقات عديدة لا حصر لها يمكن أن نوضحها بإستخدام مثل هذا النوع من الرسوم البيانية.

٢– الخطوط البيانية اللوغاريتمية Logarithmic graphs

فى حالة اذا كنا بصدد تمثيل سلسلة زمنية لبيانات تتفاوت القيم فيها تفاوتاً كبيراً أو تمثيل بيانات فى شكل معدلات النمو أو التغير السكانى من سنة لأخرى، أو معدل التغير فى الاستهلاك، أو معدلات التغير فى الدخل القومى، أو نسب تطور الدعم الحكومى للسلع والخدمات، أو نسب النقص .. والزيادة فى أى ظاهرة، فانه يجب أن يقسم المحور الرأسى إلى وحدات لوغاريتمية بدلاً من الوحدات الحسابية.

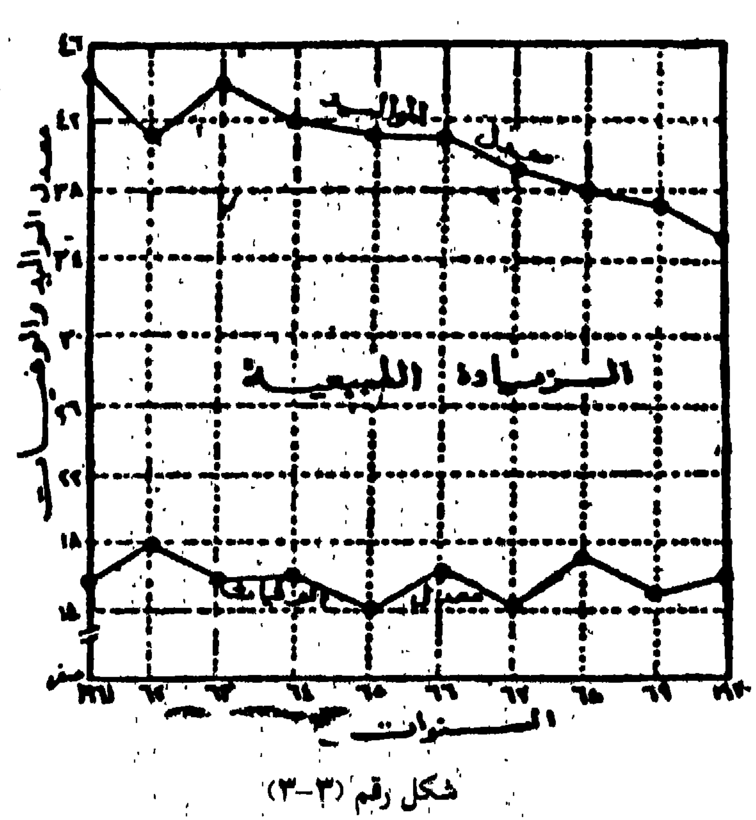
وتقوم فكرة التقسيم اللوغاريتمى على أخذ لوغاريتم الأعداد من ١ إلى ١٠ وجعلها أساساً لوحدة التقسيم اللوغارتيمى والتى تضرب فى كل مرة فى طول الدورة اللوغاريتمية المأخوذة طبقاً لطول المسافة الرأسية والأفقية المراد تمثيل الظاهرة عليها، وهى فى هذه الحالة تمثل دورة لوغاريتمية واحدة. وبعد ذلك يمكن أيضاً

أخذ دورة لوغاريتمية ثانية تبدأ بالرقم ١٠ حتى الرقم ١٠٠ وتأخذ نفس قياسات الدورة الأولى ١-١٠، كما يمكن أخذ دورة لوغاريتمية ثالثة تبدأ بالرقم ١٠٠ وتنتهي بالرقم ١٠٠٠ ولها نفس قياسات الدورة الأولى أيضاً وتمثل مئات أضعاف الوحدة الحسابية الواحدة، فاذا فرض أنه كانت لدينا بيانات أصغر رقم فيها هو ٢٠ وأكبر فيها هو ٣٠٠٠ فانه يكفي لتمثيل هذه البيانات على رسم بياني لوغاريتمي مكون من ثلاث دورات، تشمل الدورة الأولى الأقسام من ١٠ حتى ١٠٠ والثانية من ١٠٠ حتى ١٠٠٠ والثالثة من ١٠٠٠ حتى ١٠٠٠٠ ويؤخذ طول الدورة الواحدة مساويا لخمسة سنتيمترات مثلا. وقد جرت العادة على أن يبدأ التقسيم اللوغاريتمي بالرقم ١ مع قسمته على أو ضربه في الرقم ١٠ أو مضاعفاته حتى يمكن البدء بالأرقام ٠٠٠١ أو ٠٠٠٠١ كما يمكن البدء بالرقم١٠ أو ١٠٠ أو ١٠٠٠ وهكذا.



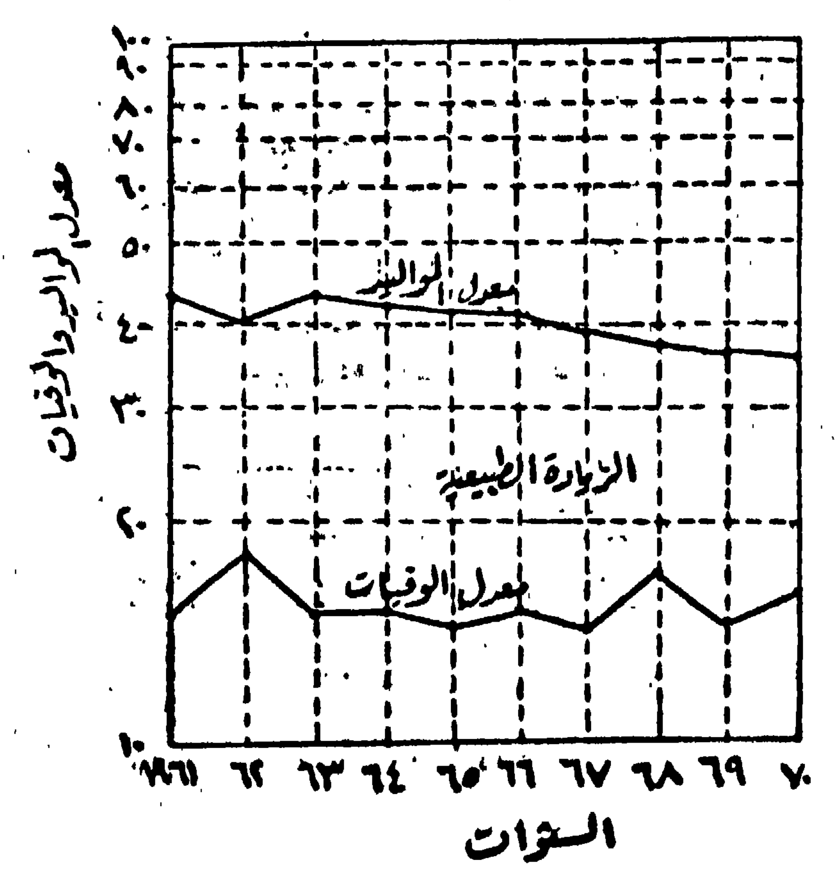
شكل رقم (٣-٢) ورق بياني لوغاريتمي مزدوج

وكما فى رسم الخطوط البيانية الحسابية وعلى حسب البيانات المتاحة يمكن أن يقسم المحور الرأسى فقط تقسيماً لوغاريتما لتوقع على أساسه معدلات التغير فى ظل الفترة الزمنية التى يقسم على أساسها المحور الأفقى تقسيماً حسابياً، يسمى ذلك بالتقسيم نصف لوغاريتمى. كما قد يقسم كل من المحورين الأفقى والرأسى تقسيماً لوغاريتمياً يطلق عليه اسم التقسيم اللوغاريتمى المزدوج، ويناسب ذلك البيانات التى تتكون من معدلات تغير أو نسب مئوية لمتغيرين مستقلين. وتبعاً لأهمية استخدام التقسيم اللوغاريتمى فى التمثيل البيانى فانه يوجد حالياً ورق رسم بيانى خاص مقسم تقسيماً لوغاريتمياً إما على المحور الأفقى أو الرأسى أو على المحورين معا (تقسيم مزدوج) شكل رقم: (٢-٣).



شكل رقم (٣-٣)
معدلات المواليد والوفيات ١٩٦١ -١٩٧٠.
(خطوط بيانية بسيطة - لاحظ كسر المحور الرأسى بعد الصفر)

والشكلان رقم (٣-٣)، (٣-٤) يوضحان مقارنة بين الخطوط البيانية الحسابية واللوغاريتمية لمعدلات المواليد والوفيات فى جمهورة مصر فى الفترة من ١٩٦١ - ١٩٧٠، ومنهما يتضح أن الخطوط البيانية اللوغاريتمية التى رسمت على رسم بيانى نصف لوغريتمى لاتظهر حدة التغير فى كل من معدلات المواليد والوفيات التى أظهرتها الخطوط البيانية الحسابية وانعكس ذلك أيضاً على الزيادة الطبيعية للسكان المحسوبة من الرسم فى كل من الشكلين.



شكل رقم (٣-٤)

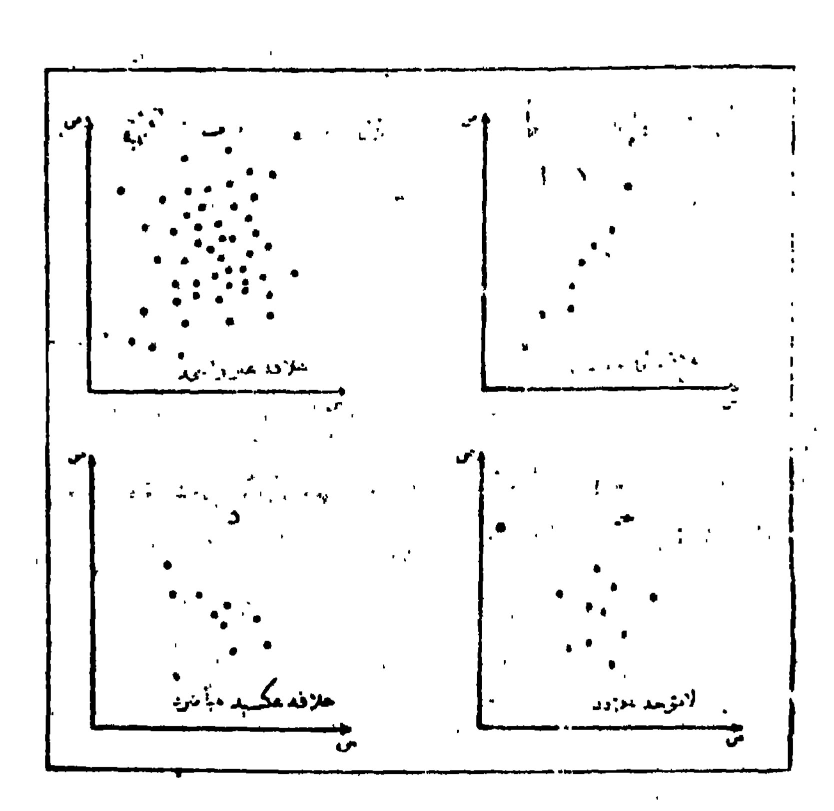
معدلات المواليد والوفيات فى الفترة من ١٩٦١ - ١٩٧٠

(رسم بيانى لوغاريتمى)

٣- أشكال الانتشار Scatter graphs

تبين أشكال الانتشار - بطريقة تقريبية - العلاقة أو مدى الترابط واتجاهاته بين متغيرين أو أكثر بفعل خصائص ومعالم معينة تربط بينها. وقد توحي هذه الأشكال بعدم وجود ارتباط بين خصائص أحد المتغيرات ومعالم المتغير أو المتغيرات الأخرى قيد البحث.

وتتمثل اتجاهات الترابط أو العلاقة بين المتغيرات فى أربعة أشكال (شكل رقم ٣-٥) : الشكل الأول منها «أ» يوضح علاقة طردية (موجبة) مباشرة بين المتغيرين، بمعنى أن تزايد قيمة أحد المتغيرين يصاحبها تزايد فى قيمة المتغير الآخر. مثال ذلك علاقة أو ارتباط أجر العمال فى أحد المصانع بكمية الوحدات المنتجة فى هذا المصنع. فإذا ارتفع أجر هؤلاء العمال ازدادت كمية الوحدات المنتجة. وبالمثل ارتباط انتاج المحاصيل الشتوية على الساحل الشمالى لمصر بكمية الأمطار، بمعنى أنه كلما ازدادت كمية الأمطار الساقطة كلما ازدهرت هذه المحاصيل وزاد انتاجها والعكس صحيح. ويتمثل الشكل الثانى (د) للعلاقة فى الارتباط العكسى أو السلبى المباشر بين المتغيرين، بمعنى أن الإختلافات أو التطور فى أحد المتغيرين يصاحبه اختلاف وتطور معاكس للتغير الآخر. مثال ذلك العلاقة الشهيرة بين كثافة السكان والبعد عن قلب المدينة. فمن المعروف أن كثافة السكان تتزايد بالقرب (أو بتناقض المسافة) من قلب المدينة والعكس صحيح. وبينما يتمثل الشكل الثالث (ب). للعلاقة بين المتغيرات فى وجود علاقة غير واضحة بين أحد المتغيرات والمتغير الآخر بمعنى أن الترابط بينهما لا يوجد له اتجاه منتظم أو محدد، فإن الشكل الرابع (جـ) للترابط يتمثل فى عدم وجود علاقة ارتباطيه على الاطلاق بين المتغيرات قيد البحث.



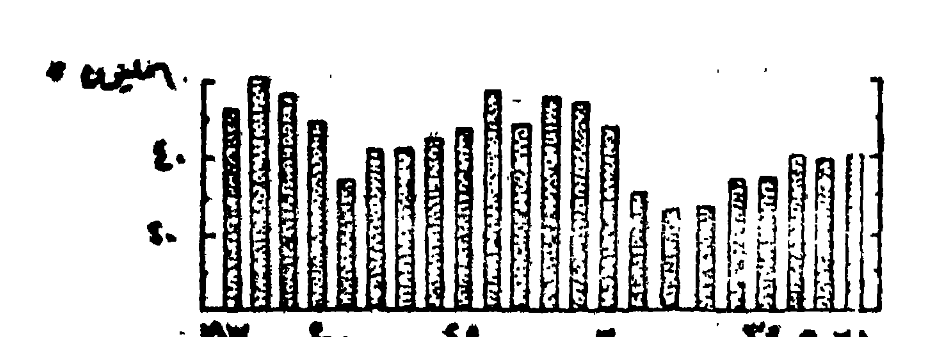
شكل رقم (٣-٥)
أشكال الانتشار لتوضيح اتجاهات الترابط بين المتغيرات

ويمكن الحصول على شكل الانتشار الذى يحدد نوع ودرجة العلاقة بين المتغيرات عن طريق توقيع بيانات المتغيرات على رسم بيانى واحد يتألف من إحداثيين أحدهما أفقى والآخر رأسى: يخصص أحدهما لتوقيع بيانات المتغير الثانى. فإذا ما وقعت القيم الممثلة للمتغيرين فى مجال انتشار متقارب أو على امتداد خط مستقيم دل ذلك على وجود علاقة قوية (طردية أو عكسية). أما إذا أظهر شكل الإنتشار تباعداً طفيفاً للقيم ولكن حول خط مستنقيم دل ذلك على وجود علاقة ضعيفة. وإذا ما سجل شكل الانتشار تباعداً كبيراً للنقط بحيث يتعذر معه أن تقع على خط مستقيم فإن ذلك يدل على عدم وجود علاقة ترابط بين المتغيرين.

٤- الأعمدة البيانية البسيطة Bar graphs

تعد طريقة الأعمدة البيانية من أيسط طرق التمثيل البيانى التى تستخدم للمقارنة بين الكميات لظاهرة واحدة أو عدة ظواهر، وعادة ماتسمى رسومها البيانية باسم Columnar diagrms. وتتألف هذه الرسوم من أعمدة ذات عرض متساوى وطول يتناسب مع الكميات التي تمثلها حسب مقياس الرسم المختار. ويمكن رسم هذه الأعمدة أما رأسيا أو أفقياً فى أشكال بيانية قائمة بذاتها، وتعتبر الأعمدة الأفقية أفضل من حيث سهولة قراءتها، أما الأعمدة الرأسية فلها ميّزة أخرى وهى سهولة المقارنة.

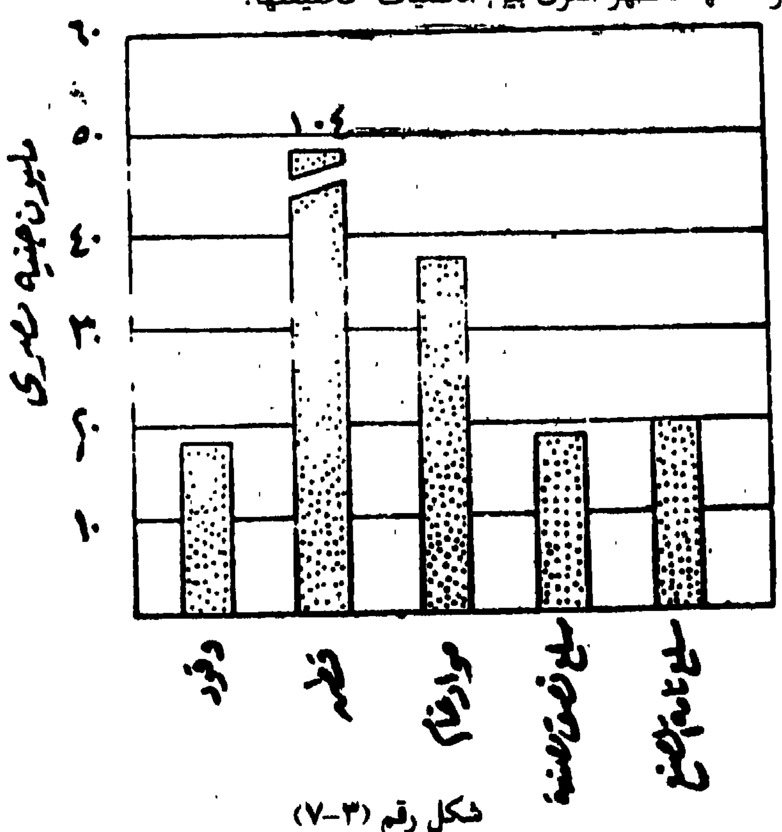
وتعتمد طريقة الأعمدة البيانية البسيطة فى تمثيل البيانات الوصفية على اظهار كميات الزيادة والنقص فى بعض الظواهر مثل معدلات المواليد والوفيات أو تمثيل التطور فى أى سلسلة زمنية خاصة بالإنتاج أو الاستثمار أو حجم المشتريات أو الدخل أو عدد السكان خلال فترات زمنية معينة. وفى هذه الحالة نرسم محورين أحدهما محور رأسى يقسم إلى أقسام متساوية تبين الكميات والآخر محور أفقى يقسم أيضاً إلى أقسام متساوية حسب الفترات الزمنية أو الصفات المميزة للظاهرة كالحالات التعليمية أو الإجتماعية أو فئات السن .. الخ ومما هو جدير بالذكر أنه عند أخذ المسافات الممثلة لقواعد الأعمدة على المحور الأفقى يجب أن تكون متساوية وعلى أبعاد متساوية أيضاً، وذلك بطريقة تلائم المساحة من لوحة الرسم المخصصة للتمثيل البيانى وعدد الأعمدة المراد رسمها (شكل رقم ٣-٦). وفى حالة الفترات الزمنية غير المنتظمة فإن المسافات بين كل عمود وآخر يجب أن يتناسب مع الأبعاد الزمنية للفترة المراد تمثيلها بيانياً. كما يجب فى كل الحالات أن يبدأ المقياس على المحور الرأسى من الصفر وينتهى برقم أعلى من أكبر قيمة من قيم الظاهرة موضع البحث. إلا أنه فى كثير من الحالات نجد بين قيم الظاهرة المراد تمثيلها بطريقة الأعمدة البسيطة قيمة أو قيمتين متطرفين أو شاذتين تفوق بقية قيم الظاهرة بما يؤدى إلى وجود تفاوت كبير لهذه القيم. وبالتالى يؤدى ذلك إلى



شكل رقم (٣-٦)
قيمة منتجات مصايد الأسماك الكندية
في الفترة ١٩١٧ - ١٩٣٨ «بالمليون دولار»

اختلاف كبير في طول الأعمدة، بل أنه في بعض الأحيان يصبح من الصعب تمثيل القيم بأخذ مقياس رسم على المحور الرأسي يلائم هذه القيم المتفاوتة. فمثلاً إذا كانت لدينا كمية أكبر مائة مرة من كمية أخرى، فأنها تتطلب رسم عمود أطول مائة مرة من عمود الكمية الأصغر. وهذا يضطرنا طولها الكبير الصغيرة بأعمدة صغيرة جداً، وإما أن نرسم أعمدة قد يضطرنا طولها الكبير جداً إلى تقطيعها إلى قطاعات توضع بجوار بعضها البعض. ولو أن كل هذا التحايل لاينقل الصورة الصحيحة لتمثيل هذه الكميات، ومالذلك من تقليل من أهمية هذا الأسلوب في التمثيل البياني. وللتغلب على هذه المشكلة يستحسن قطع المحور الرأسي الموجود عليه لقياس الكميات وجعل الجزء الأسفل منه يبدأ من الصفر حتى قيمة أعلى من الكمية الصغيرة، أما الجزء الأعلى فيبدأ من رقم أقل من الكمية الكبيرة وينتهي برقم أعلى منها مع ثبات طول المقياس في الجزئين. وفي بعض الأحيان تكسر الأعمدة التي تمثل قيما متطرفة ويكون ذلك بالتخلص من

الارتفاعات العادية للقيم الأخرى وقد يكون كسر الأعمدة رأسياً عن طريق وضع خطين متوازيين مائلين عند نهاية الارتفاع المراد تحديده والذى يناسب الشكل وليدل على أن العمود بقية، ولكن مساحة ورقة الرسم لاتسمح باظهارها، ويجب أن نكتب أعلى هذا العمود بالذات الكمية الحقيقية التى يمثلها (شكل رقم ٣-٧). وعلى الرغم من ذلك فان هذه الطريقة لا يمكن الاستفادة بها فى حالة المقارنة لأنها لاتظهر الفرق بين الكميات كحقيقتها.



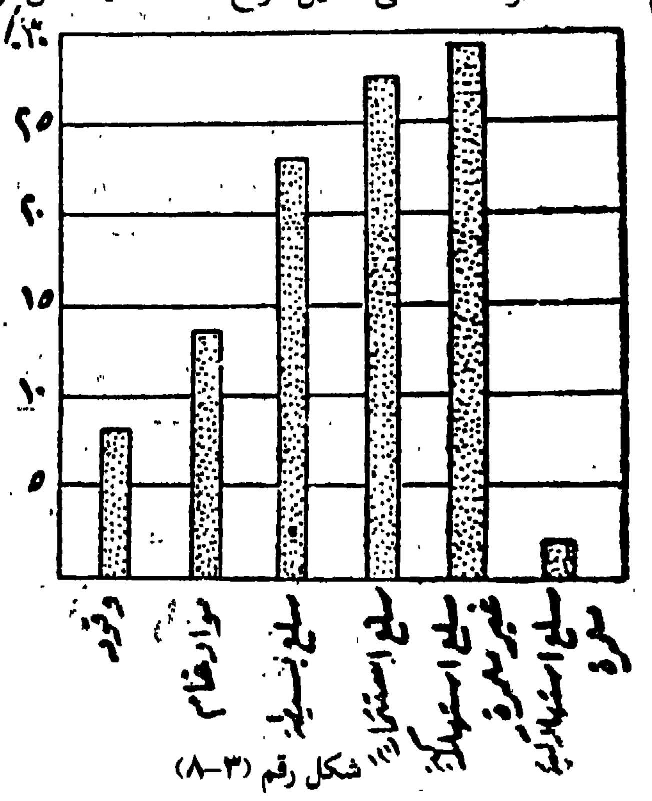
شكل رقم (٣-٧)

قيمة الصادرات المصرية فى عام ٦٢، ١٩٦٣

بالأعمدة البيانية البسيطة (لاحظ كسر العمود الثانى)

وهناك نوع آخر من الأعمدة البيانية يصلح فى اظهار الأهمية النسبية لكمية الظاهرة يسمى بالأعمدة النسبية Proportional Bargraph والتى ترسم لتمثيل أعداد السكان أو الإنتاج المعدنى أو حركة الصادرات والواردات فى الموانئ. وتتميز

طريقة الأعمدة النسبية بسهولة رسمها من ناحية التصميم، وكذلك بسهولة القراءة من الناحية المرئية. وتتلخص طريقة رسم الأعمدة النسبية فى أن نبدأ أولاً باستخراج النسبة المئوية للكميات التى نريد تمثيلها بالنسبة للمجموع الكلى للكميات، مثل نسبة وزن المجموعات الرئيسية للواردات المصرية فى سنة ١٩٦٢، ويخصص المحور الأفقى لتعيين المجموعات المختلفة للواردات، أما المحور الرأسى فيخصص للأوزان المناظرة لكل مجموعة ويقسم إلى أقسام متساوية تبين النسبة المئوية للأوزان مبتدئين بالصفر ومنتهين بنسبة أعلى من أعلى النسب المراد تمثيلها كما فى الشكل رقم (٣-٨). ويحسن عند رسم هذا النوع من الأعمدة أن نختار له نوع من التظليل المصمت (كاللون الأسود المصمت) أو نستخدم نمط التظليل النقطى وذلك لأن استخدام نمط الخطوط المائلة فى تظليل فراغ الأعمدة يتضمن نوعاً من خداع البصر.



شكل رقم (٣-٨)

أوزان المجموعات الرئيسية للواردات المصرية ٦٢/ ١٩٦٣

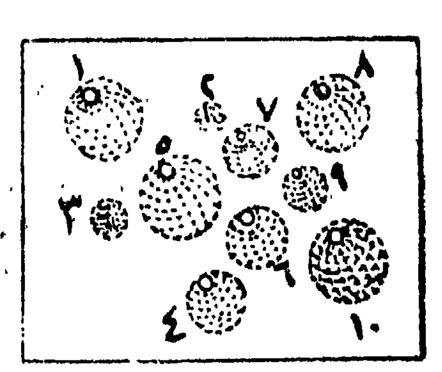
الأعمدة البيانية النسبية

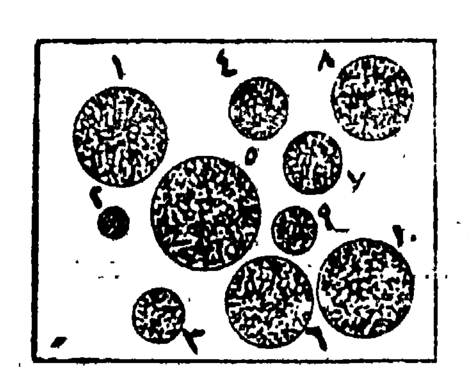
٥- الرسوم البيانية الحجمية Three- dimensional graphs

إذا كانت البيانات المراد تمثيلها بيانياً ذات مدى عظيم جداً في القيم أو الكميات، فاننا ندخل البعد الثالث الذى يترتب عليه استخدام رسوم بيانية حجمية تتناسب أحجامها مع مقدار الكميات التى تمثلها. ومن أهم هذه الرسوم البيانية الكرات Spheres والمكعبات Cubes. وعلى الرغم من مميزات هذا النوع من الرسوم البيانية فان هناك بعض من المثالب التى يمكن إجمالها فى: أن رسم الرسوم الحجمية ليس أمراً سهلاً بل يتطلب جهداً وعملاً إضافياً حتىّ يبدو الشكل الحجمى واضحاً، أو بمعنى آخر أن نعطى الكرة أو المكعب الشكل الحجمى بأبعاده الثلاثة على سطه لوحة الرسم المستوى. وعلى الرغم من أن العلاقة بين أحجام الأشكال والكميات التى تمثلها صحيحة رياضية إلا أنه ليس من السهل تقدير أحجام هذه الأشكال بمجرد النظر إليها عكس الأعمدة البيانية.

وفى حالة إستخدام الرسوم البيانية الحجمية لتمثيل ومقارنة كميات عظيمة التفاوت والاختلاف فان حجم هذه الأشكال تتناسب مع مكعب نصف القطر (فى حالة الكرات) أو مع مكعب طول الضلع (فى حالة المكعبات) فالكرة الأكبر عشرة مرات من كرة أخرى سوف تمثل كمية أكبر ألف مرة (١٠^{٣}) من الكمية التى تمثلها الكرة الأخرى. وكما هو متبع فى طريقة رسم الدوائر البيانية، فاننا نستخرج أولا الجذور التكعيبية للكميات، ونعتبر هذه الجذور التكعيبية أنصاف أقطار للدوائر التى سنعطيها شكل الكرات، أو نعتبرها طول المكعبات المراد رسمها. وفى حالة رسم الكرة نبدأ أولاً برسم دائرة عادية ثم نعطيها الشكل الحجمى، أما أن نجعلها تمثل شكل «الكرة الأرضية» وذلك برسم شبكة رمزية من دوائر العرض وخطوط الطول فوق الدائرة المفرغة والتى ستبدو فى النهاية على شكل كرة مجسمة، وأما أن نطمس كل مساحة الدائرة باللون الأسود مع ترك مساحة بيضاء فى أعلى الكرة بحيث تبدو كالنور الساطع فى أعلى الكرة كما فى الشكل رقم (٣-٩ب)، أما المكعبات فهناك نوعان منها: نوع يبدو على شكل الدولاب وفيه يكون طول الجوانب مساوية لنصف طول الوجهة، والنوع الآخر يبدو متساوى الأضلاع والارتفاعات وتكون أطوال الجوانب والوجهة متساوية. ويعد أحسن شكل للمكعب

هو الذى يكون فيه طول ضلع جوانبه ٣/٥ طول ضلع واجهته، بحيث تميل هذه الجوانب من ٣٠ إلى ٥٠ من الخط الأفقى وتكون جوانب المكعب على يمين الناظر للرسم البيانى.

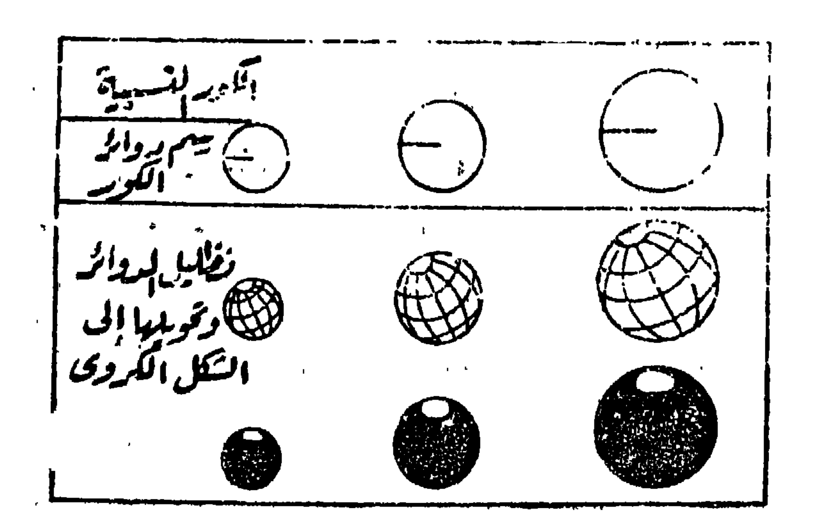




شكل رقم (٣-٩أ)

الرسوم البيانية الحجمية «كرات» والمساحية «دوائر»

لاحظ اختلاف حجم الكرات عن مساحة الدوائر لنفس البيانات



شكل رقم (٣ - ٩ب)

الكرات النسبية بيانات سكان المدن المصرية لعام ١٩٦٦

وكمثال لتطبيق طريقة الرسوم الحجمية يمكن تمثيل عدد سكان كل من القاهرة والاسكندرية والجيزة (أكبر المدن المصرية) بالكرات أو المكعبات كما هى الحال فى الشكل رقم (٣- ٩ب) الذى يعتمد على البيانات للجدول التالى:

جدول رقم (٣-١)

عدد سكان أكبر المدن المصرية عام ١٩٦٦

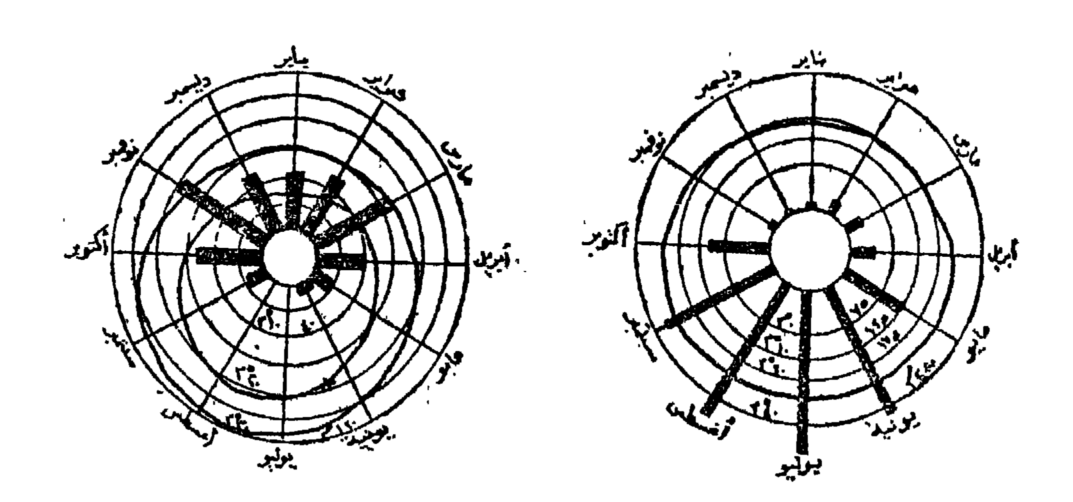
عدد السكان المدينة	عدد السكان (بالآلاف)	الجذر التكعيبى	طول قطر الكرة أو ضلع المكعب (الجذور التكعيبى / ٢٠٠)
القاهرة	٤,٢٢٠	١٦١,٦٠	٠,٨ سنتيمتر
الاسكندرية	١,٨٠١	١٢١,٦٤	٠,٦ سنتيمتر
الجيزة	٥٧٠	٨٢،٩٧	٠,٤ سنتيمتر

٦- الرسوم البيانية الدائرية Circular (Colck) graphs

تستخدم طريقة الرسوم البيانية الدائرية (Polar Chart) لتمثيل بيانات المتغيرات المستمرة، مثل بيانات درجة الحرارة والمطر لمختلف محطات الأرصاد الجوية فى بلدما، أو توزيع متوسط نسبة العاملين فى أقسام أحد المحلات التجارية الكبرى على شهور السنة، أو توزيع العمالة فى مختلف النشاط الزراعى على مدار شهور السنة. وتتفوق هذه الطريقة على طريقة الخطوط البيانية البسيطة، فى أن نهايات الرسم البيانى (الطرف الأيمن والطرف الأيسر) فى الطريقتين الأخيرتين يقطع استمرار تطور الظاهرة أو يخفى اتجاهها العام على مدى فترة التوزيع.

وتعتمد طريقة الرسوم الدائرية على رسم دائرة وتقسيمها إلى ١٢ قسماً تبعاً لعدد شهور السنة بحيث يكون نصيب كل شهر ٣٠ درجة من درجات الدائرة (٣٦٠ درجة)، ثم يرسم من مركز هذه الدائرة كميات المتغير قيد البحث، مثل درجات الحرارة المختلفة، على أن نراعى أن يشمل ذلك أقل الكميات (الدرجات) وأعلاها. فإذا كان مقدار أو كمية المتغير لاتقل عن الصفر ففى هذه الحالة ج

يصبح مركز الدائرة هو الصفر، أما إذا كان هذا المقدار أقل من الصفر، فيصبح مركز الدائرة في هذه الحالة ممثلا لرقم يقل عن الصفر فى حين يمثل رقم الصفر نفسه بدائرة منفصلة. فمثلاً إذا كانت لدينا بيانات عن المتوسطات الشهرية لدرجة الحرارة وكمية المطر لمحطات جوية تتميز بفصلية توزيه هذين العنصرين المناخيين فإنه يمكن تمثيل هذه البيانات برسم بيانى دائرى عن طريق رسم أنصاف أقطار على بعد من مركز الدائرة الرئيسية «الخارجية» يمثل كل منها بيانات كل عنصر تبعاً لقيمتها. ثم نصل بين نهايات هذه الأنصاف أقطار بخط منحنى وبذلك تتكون لدينا شبه دائرة محيطها متعرج يمثل عنصر درجة الحرارة كما نرسم على أنصاف الأقطار أعمدة بيانية بسيطة تمثل أطوالها كميات المطر الساقطة فى كل شهر من شهور السنة (شكل رقم ٣-١٠).



شكل رقم (٣-١٠)

الرسوم بيانية الدائرية لتمثيل عنصرى

درجة الحرارة والمطر فى كلكتا «الهند» وملقا «أسبانيا»

ويمكن تقسيم شبه الدائرة الناتجة عن تمثيل درجة الحرارة إلى شرائح تمثل كل شريحة منها فصلاً حرارياً مميزاً مثل الفصل الحار والفصل البارد أو الفصل الذي تسمح فيه درجة الحرارة بنمو النباتات مثلاً. كما يجوز عن طريق هذا الرسم تحديد اليوم الذى يبدأ فيه كل فصل من هذه الفصول الحرارية.

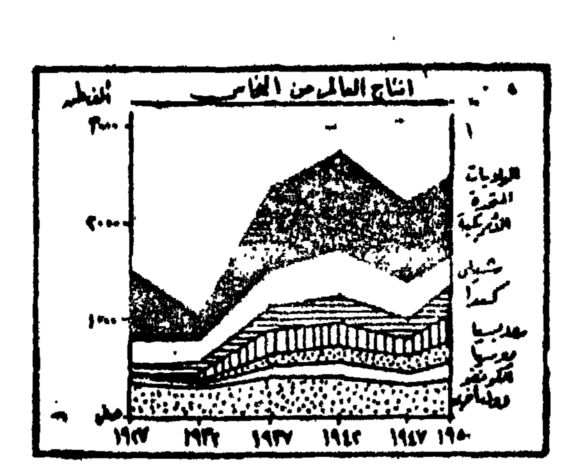
وعلى الرغم من تغلب طريقة الرسوم البيانية الدائرية على مشكلة انفصام توزيع البيانات المستمرة على الرسوم البيانية العادية «طريقة الخط البيانى البسيط مثلاً»، إلا أن استخدامها ليس شائعاً أو متداولاً مثل بقية الطرق البيانية بسبب أن تذبذب «ارتفاع وانخفاض» الخط البيانى الذى يمثل بيانات المتغير بمكن ملاحظته مباشرة على الرسم البيانى العادى، بينما يظهر ذلك فى الرسم الدائرى على شكل إقتراب أو ابتعاد عن مركز الدائرة، مما يتطلب دقة ومهارة خاصة عند قراءة وتفسير الشكل الناتج عنها.

ثانياً: الطرق البيانية لتمثيل التغير فى مكونات الظاهرة والمجموع الكلى لها:

إن الهدف الأساسى من هذه المجموعة من الطرق البيانية هو المقارنة بين قيم المكونات المختلفة لظاهرة ما بعضها البعض وبين كل منها والمجموع الكلى للظاهرة. وفيمايلى دراسة مختصرة توضح غرض وأسلوب إنشاء كل طريقة من الطرق الشائعة الاستخدام فى مجال العرض البيانى للبيانات الاحصائية وهى: الرسوم البيانية المجمعة، الأعمدة البيانية المركبة، الأهرامات البيانية، الدوائر المقسمة الرسوم المثلثية.

١ - الرسوم البيانية (المنحنيات) المجمعة Compound- line graphs

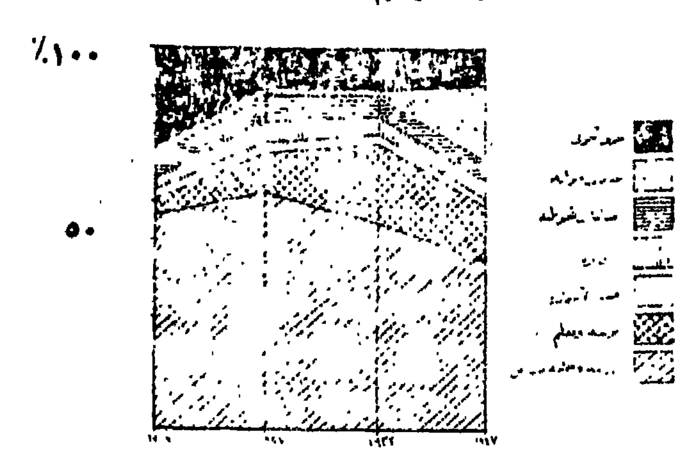
وهى عبارة عن خطوط بيانية تمثل التغير فى مجموع الظاهرة الواحدة على مدى فترة زمنية أو التغير فى مجموع الظاهرة وظاهرة (أو ظاهرات) أخرى، بحيث يمثل التغير فى أجزاء أو مكونات الظاهرة بخطوط بيانية بسيطة يمثل كل خط منها بداية القياس الخط اللاحق له، ثم تظلل المساحات المحصورة بين هذه الخطوط البيانية (شكل رقم ٣-١١)



شكل رقم (٣-١١)

إنتاج النحاس فى العالم «بالآلاف الأطنان المترية»

«طريقة الرسوم البيانية المجمعة»



شكل رقم (٣-١٢)

النسب المئوية للمشتغلين بكل من الحرف الرئيسية

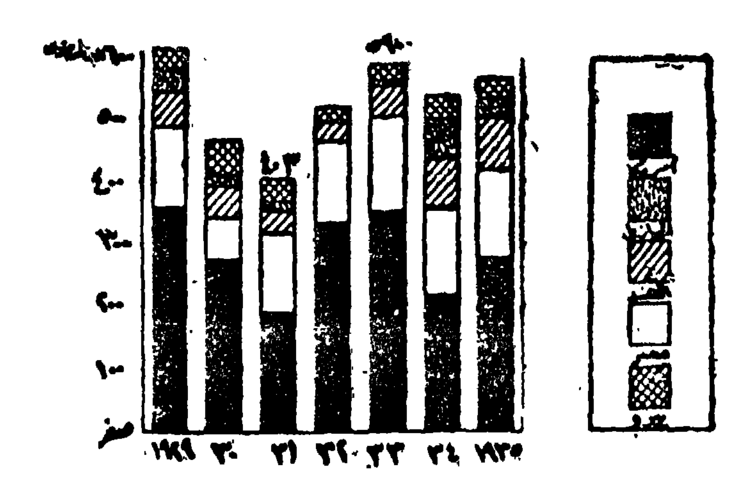
فى أحد المحافظات فى الفترة من ١٩١٧ - ١٩٤٧

«طرق الرسوم البيانية المجمعة النسبية»

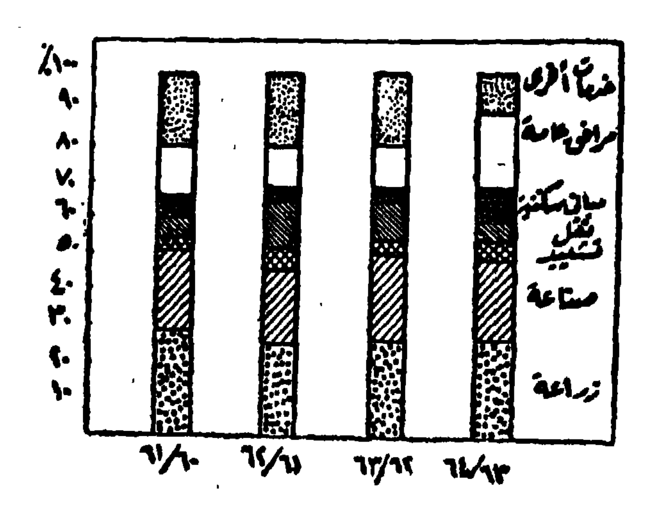
ويمكن رسم هذا النوع من الرسوم على أساس النسب المئوية هى بطبيعة الحال تكون الأنسب والأحسن . ويتم ذلك بتقسيم الظاهرة إلى أجزائها المختلفة بشرط أن تكون بنفس الترتيب لكل فترة زمنية، ثم نصل بين نقط التقسيم بخطوط وتظلل المساحات المحصورة بين هذه الخطوط وبذلك يمكن معرفة عما إذا كانت نسبة أى قسم من الظاهرة قد هبطت أو زادت فى نفس الوقت بالنسبة إلى باقى التقسيمات الفرعية الأخرى للظاهرة (شكل رقم ٣-١٢).

٢- الأعمدة البيانية المركبة Compound- Bar graphs

أما طريقة الأعمدة البيانية المركبة فهى عبارة عن أعمدة ذات عرض متساوى ومقسمة إلى أقسام داخلية تمثل فى مجموعها المجموع الكلى للظاهرة، وذلك بدلا من رسم عدة رسوم بيانية كل رسم منها يمثل جزءاً أو قسماً من المجموع الكلى للظاهرة. وفى هذه الحالة فانه يمكن مقارنة كميات هذه الظاهرة من ناحية الكميات المطلقة -شكل رقم ٣-١٣)- كما أنه يمكن مقارنة كميات هذه الظاهرة من الناحية النسبية وذلك بتحويل كمية كل قسم فرعى منها إلى نسبة مئوية. وتسمى الأعمدة البيانية فى هذه الحالة باسم الأعمدة المركبة النسبية (شكل رقم ٣-١٤). وفى هذه الحالة لا يمكن مقارنة كل عمود (مستطيل) بآخر ولكن يمكن مقارنة الجزئيات (التفاصيل) من كل عمود بالجزء الذى يناظره فى العمود الآخر، وذلك بمعرفة الفرق بين نسبيتهما بالنسبة للمجموع الكلى. ويجب فى هذه الحالة أن يصاحب الرسم البيانى للأعمدة المركبة النسبية. رسم بيانى آخر تكون أرقامه مطلقة حتى يمكن معرفة التغير بين المجموع الكلى لكل ظاهرة وأخرى.



شكل رقم (٣-١٣)
الواردات البريطانية من الأقطان الخام
«طريقة الأعمدة المركبة المطلقة»



شكل رقم (٣-١٤)
التطور النسبى لمكونات الدخل القومى لمصر
فى الفترة ٦٠- ١٩٦٤ «طريقة الأعمدة المركبة النسبية»

٣- الأهرامات البيانية Pyramid graphs

تستخدم الأهرامات البيانية كأحد طرق التمثيل البياني للبيانات الديموجرافية وبصفة خاصة لبيانات التركيب النوعي والعمرى للسكان، حيث يجمع الهرم البياني نسب كل من الذكور والإناث إلى العدد الكلي للسكان فى الفئات العمرية المختلفة. والهرم البيانى عبارة عن أعمدة بيانية أفقية ترسم على محورين أفقين أحدهما يمثل أعداد (أو نسب) السكان الذكور والآخر يمثل أعداد (أو نسب) السكان الإناث، أما المحور الرأسى للهرم فهو يمثل فئات العمر لكل من النوعين من السكان ويجب أن تقسم المحاور الأفقية بنفس المقياس سواء الذكور أو الإناث.

ومن المفيد إستخدام هذا الأسلوب من التمثيل البيانى إذا أريد معرفة الخصائص أو تشخيص الاتجاهات للمجتمعات السكانية، وكذلك عند إجراء المقارنات بين حالة السكان لأكثر من إقليم أو دولة لإظهار الصفات العامة للسكان باستخدام بيانات التعدادات السكانية. وهناك عدة أنواع من أهرامات السكان البيانية نوجز طريقة أنشاء كل منها فيمايلى:

أ- الهرم السكاني البسيط:

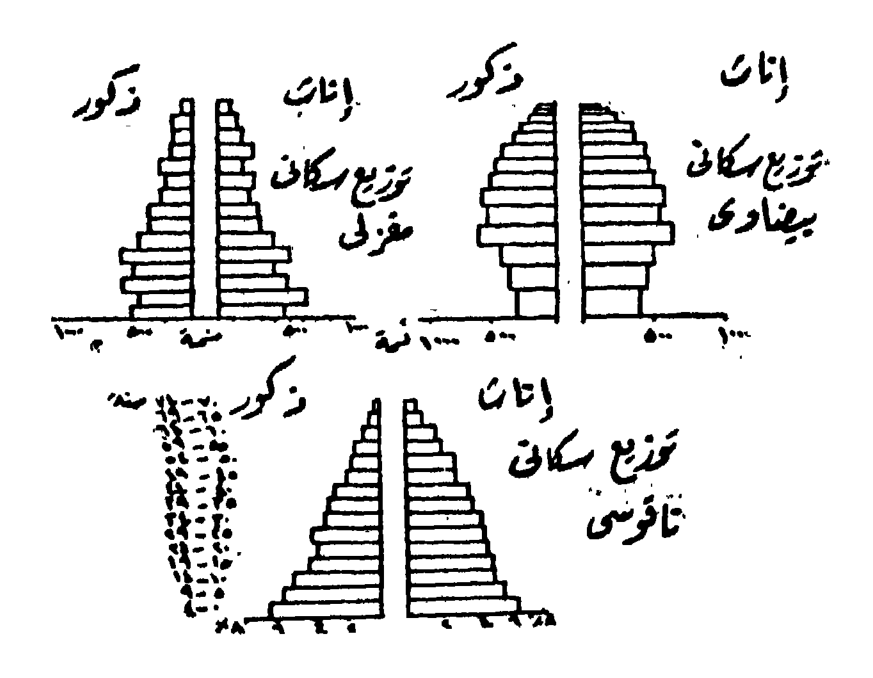
وتقوم فكرة إنشائه على الأساس السابق شرحه لإنشاء ورسم الهرم السكاني. ويستخدم هذا النوع من الأهرامات لبيان الصفات العامة لسكان دولة أو أقليم معين. ومن المعروف لدى علماء الديموجرافيا أن لكل دولة هرم سكاني يميز تراكيبها السكاني من حيث النوع والعمر لتعداد معين. وبناء على ذلك فإن أشكال الأهرامات السكانية ستختلف باختلاف التركيب النوعي والعمرى للسكان بين البلاد المختلفة وهذا الإختلاف هو الذى يبرز المميزات ويؤكد الاتجاهات السكانية بين البلاد المختلفة. وهذا الإختلاف هو الذى يبرز المميزات ويؤكد الإتجاهات السكانية التى بالتالى تعطى صورة واضحة عن التركيب العمرى والنوعي للمجتمعات السكانية لهذه الدول (شكل رقم: ٣-١٥) فمثلاً إذا كان الهرم السكانى يأخذ الشكل المغزلى المقلوب فان ذلك يدل على أن المجتمع الذى يمثله يتميز بتعادل معدلات المواليد والوفيات فيه. وإذا كان الهرم السكاني يأخذ

الشكل البيضاوى من أعلى (أى فى الفئات ذات الأعمار الكبيرة) وليس من عند القاعدة (أي في الفئات ذات الأعمار الكبيرة) وليس من عند القاعدة (أى فى الفئات ذات الأعمار الكبيرة) وليس من عند القاعدة (أى فى الفئات ذات الأعمار الصغيرة) فانه يدل على أن المجتمع الذى يمثله مجتمعا مسنا. ويستنتج ذلك من إنخفاض نسبة الأطفال (ذكور وإناث) وزيادة نسبة المسنين (الذكور والإناث). أما إذا كان الهرم السكانى يتخذ شكلا قريبا من شكل الناقوس (الجرس) حيث تكون قاعدته عريضة ومحدب بلطف فى قمته فإن ذلك يدل على ارتفاع معدلات الخصوبة.

ب- الاهرامات السكانية المركبة Compound Pyramids

وتقوم فكرة هذا النوع من الأهرامات السكانية على أساس تمثيل التركيب النوعى أو العمرى للسكان بأعمدة طول كل عمود يتناسب مع العدد الكلى للسكان لكل تعداد من التعدادات وبعد ذلك يقسم العمود (مثل طريقة الأعمدة المركبة) إلى أقسام الظاهرة الفرعية، كأن يقسم مثلا إلى سكان الريف (الزراع وغير الزراع) وسكان الحضر لكل تعداد. ففى الشكل رقم (٣-١٦) يمكن ملاحظة أنه خلال الفترة من ١٧٩٠ إلى ١٨٨٠ كان هناك فئتين فقط (سكان الريف والمحلات العمرانية التى يتراوح عدد السكان بها من ٢٥٠٠ إلى ٨٠٠٠ نسمة، بالإضافة إلى فئة السكان أكثر من ٨٠٠٠ نسمة والتى يمكن أن نعتبرها ممثلة لسكان الحضر. أما فى الفترة من ١٨٩٠ إلى ١٩١٠ فإننا نلاحظ ثلاث تقسيمات عندما أدخلت فئة السكان ٢٥٠٠ - ٨٠٠٠ نسمة كفئة مستقلة.

ويلاحظ على أعمدة الهرم من سنة ١٩٢٠ حتى ١٩٤٠ أنه قد ارتفع عدد تقسيماتها الداخلية لتضم فئتين لسكان المناطق الريفية من الزراع وغير الزراع، بالإضافة إلى فئتين لتضم فئتين أخريتين: فئة السكان الذين يتراوح عددهم من ٢٥٠٠ إلى ٨٠٠٠ نسمة، وأخرى للسكان أكثر من ٨٠٠٠ نسمة (سكان الحضر).



شكل رقم (٣-١٥)
أشكال الأهرامات السكانية البسيطة

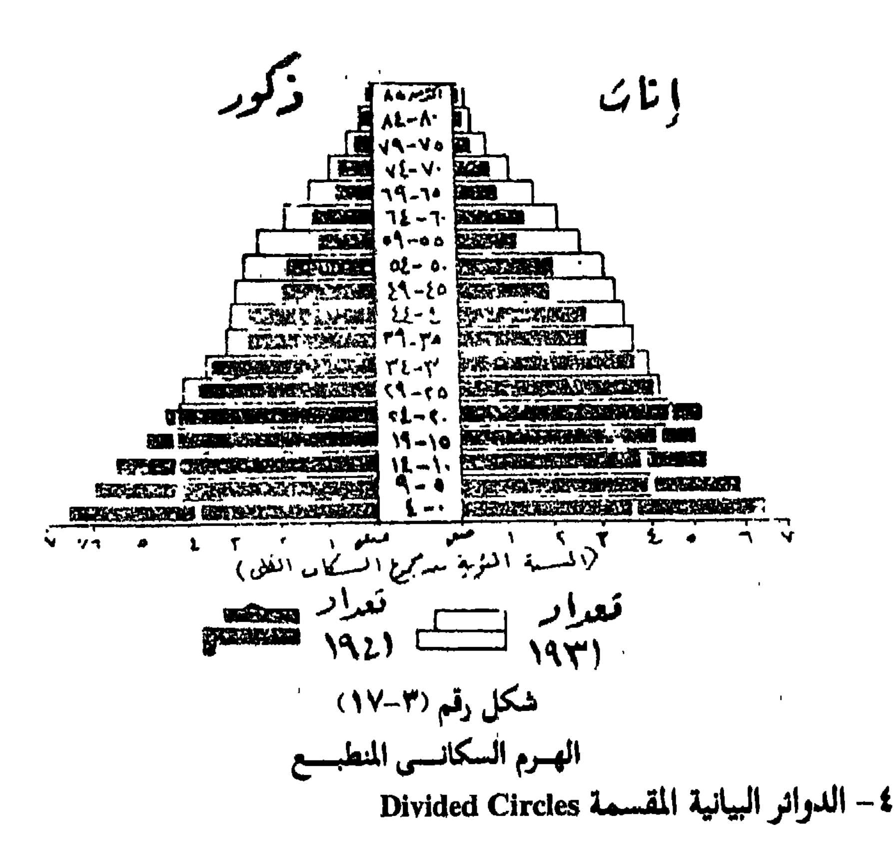
جـ- الاهرامات السكانية المنطبعة Superimposed Pyrmids

ويستخدم هذا النوع من الأهرامات لتمثيل بيانات التركيب النوعي والعمرى فى مكان ما لعدة تعدادات مختلفة وذلك بقصد المقارنة بين عدد السكان فى كل تعداد وآخر. كما يمكن إستخدام هذه الطريقة لمقارنة حالة السكان من حيث التركيب العمرى والنوعى لمنطقتين للوقوف على مدى إختلاف توزيع السكان فى إحداهما عن الأخرى.

شكل رقم (٣-١٦)

الهـــرم البيانـــى المركـــــب

وفى كل من الحالتين يمكن رسم هرم سكانى بسيط لأحد التعدادين أو لإحدى المنطقتين بالطريقة السابق ذكرها وإعطائه لونا أو ظلا معينا. ثم يرسم بعد ذلك هرماً بسيط اخر للسكان فى التعداد الثانى أو فى المنطقة الثانية بنفس مقاييس الرسم المستخدمة على الهرم السكانى الأول فيبدو وكأنه منطبعاً عليه (شكل رقم : ٣ - ١٧).



شكل رقم (٣-١٧)

الهـرم السكانـــى المنطبـــع

٤- الدوائر البيانية المقسمة Divided Circles

يهدف رسم الدوائر البيانية المقسمة، التى تستخدم أساساً لمقارنة ظاهرتين أو أكثر، أو للمقارنة بين المكونات المختلفة لظاهرة واحدة بعضها البعض والمجموع الكلى خلال فترات زمنية متفاوتة، إلى إظهار التفاوت بين مجموع قيم الظاهرات وبيان التغيرين المجموع الكمى لقيم الظاهرة الواحدة ومكوناتها. ويمكن أن نستفيد من إستخدام هذه الدوائر فى حالتين أساسيتين:

عندما يكون المجموع الكلى كبيراً نسبياً ولكنه يتمثل فى مساحة محدودة جداً، عندما يكون المجموع الكلى كبيراً نسبياً ولكنه يتمثل فى مساحة محدودة جداً، كما فى حالة تمثيل عدد سكان المدن أو تمثيل إنتاج المصانع، أو عندما نريد تمثيل الكميات الكلية فى منطقة أو أقليم أو دولة، كما فى حالة تمثيل إنتاج البترول فى البلاد العربية مثلا.

ونظراً لأن مساحة الدائرة تتناسب مع مربع نصف قطرها (مساحة الدائرة = ط نق٢) فانه يجب عند رسم مجموعة من الدوائر البيانية لعدة ظاهرات أخذ للجذر التربيعي للقيم الكلية التي تمثل كل ظاهرة ليكون بمثابة نصف قطر لكل دائرة على حدة. أما إذا كان التمثيل البياني لظاهرة واحدة فقط فعلينا أن نختار طول نصف القطر المناسب والذى يعطى مساحة لدائرة تتلائم مع مساحة ورقة الرسم المراد تمثيل الظاهرة عليها.

ولتمثيل البيانات الموضحة فى الجدول رقم (٣-٢) بطريقة الدوائر البيانية تجرى الخطوات الآتية:

أ - نجمع الإنتاج فى كل سنة حتى نحصل على المجموع الكلى (السنوى) للإنتاج فى كل فترة زمنية.

ب- نستخرج الجذر التربيعي لمجموع الإنتاج فى كل سنة على حدة، ويكون الناتج ممثلاً لطول لنصف القطر (نق) الذى يريد أن نعرفه لكى ترسم الدوائر التى تمثل الإنتاج، والجذور التربيعية للشال هى ١١,٣ ، ١١,٦ ، ١٢.

جدول رقم (٣-٢)

إنتاج مناطق الصيد بجمهورية مصر فى الفترة ٦١ - ١٩٦٤ (بالألف طن)

الإنتاج / السنوات	١٩٩٢/٦١	١٩٩٣٤/٦٢	١٩٩٦٤/٦٣
إنتاج البحار	٦٥	٦٥	٧٠
إنتاج البحيرات	٤٧	٥٤	٥٦
إنتاج النيل	١٥	١٦	١٨
المجموع	١٢٧	١٣٥	١٤٤

جـ- نختار قيمة قياسية أساسية سواء بالسنتيمتر أو الملليمتر، يمكن على أساسها أن نحول أعداد الجذور التربيعية الناتجة لدينا إلى أطوال متناسبة تمثل مباشرة أنصاف أقطار الدوائر. وفى العادة تعطى هذه القيمة الأساسية لأصغر جذر تربيعى.

د- ولمعرفة أنصاف أقطار الدوائر، هناك عدة طرق هامة تؤدى كلها إلى نتيجة واحدة ولمنها تختلف فى العمليات الحسابية. وسنختار من هذه الطرق طريقتين مألوفتين هما: طريقة التناسب الحسابى (طريقة المقص). ويمكن أن نطبقها على المثال السابق. فمثلاً إذا اخترنا الطول ١٦ ملليمتر كقيمة أساسية للجذر التربيعى ١١,٣ فان:

(١) ١١,٣ = ١٦ ملليمتر (نصف قطر الدائرة الأولى)

١١,٦ = س ملليمتر (نسف قطر الدائرة الثانية)

$$س = \frac{١١,٦ \times ١١٦}{١١,٣} = ١٦,٤ \text{ ملليمترا}$$

(٢) ١١,٣ = ١٦ ملليمتر (نصف قطر الدائرة الأولى)

١٢ = س ملليمتر (نصف قطر الدائرة الثالثة)

$$س = \frac{١٦ \times ١٢}{١١,٣} = ١٧ \text{ ملليمترا}$$

والقيمة الأساسية التى اخترناها يعتمد اختيارها على مساحة لوحة الرسم ويجب عند اختيارها أن تتوافق مساحة الدوائر مع أبعاد مسطح لوحة الرسم بحيث لا تظهر أصغر دائرة صغيرة جداً وأكبر دائرة كبيرة جداً بالنسبة لمساحة لوحة الرسم.

أما الطريقة الأخرى فتعرف بطريقة الجذور التربيعية وهى طريقة سهلة ولا تتطلب كثيراً من الحساب وتتلخص فى أن نقسم الجذور التربيعية على العدد ١٠ أو قوى هذا العدد الصحيح (١٠٠، ١٠٠٠، ١٠٠٠٠ ... الخ) وذلك طبعاً على

حسب المدى الذى توجد عليه الجذور التربيعية. ففى مثالنا السابق يمكن أن نقسم الجذور التربيعية كلها على العدد ١٠ ويكون تمييز الناتج بالسنتيمتر، وعلى هذا الأساس نجد أن:

نصف قطر الدائرة الأولى= ١١,٣ ÷ ١٠ = ١,١٣ سنتميتراً
نصف قطر الدائرة الثانية = ١١,٦ ÷ ١٠ = ١,١٦ سنتيمتراً
نصف قطر الدائرة الثالثة = ١٢ ÷ ١٠ = ١,٢ سنتميتراً

وكما هو واضح فان الأطوال التى نتجت بهذه الطريقة هى أطوال صغيرة وبالتالى ستكون مساحات دوائرها صغيرة. وفى مثل هذه الحالة يجب أن نكبر الأطوال الناتجة، وذلك بضربها كلها فى أى رقم نختاره، بحيث تظهر الدوائر بعد رسمها ملائمة لأبعاد لوحة الرسم. فإذا كان هذا الرقم الذى اخترناه هو ١,٥ مثلا فسوف يصبح:

طول نصف قطر الدائرة الأولى= ١,١٣ × ١,٥ = ١,٦٩٥ سنتميتراً (١٫٧٠ سم تقريبا)

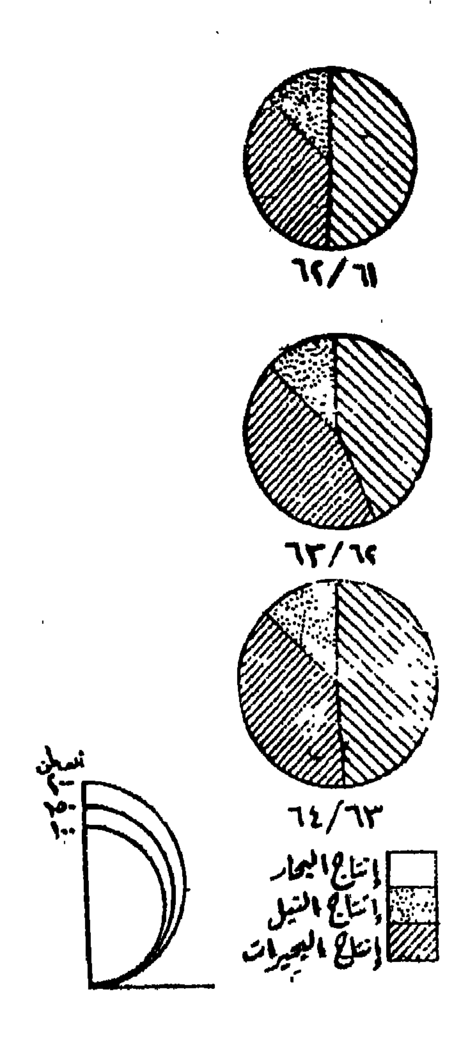
طول نصف قطر الدائرة الثانية = ١,١٦ × ١,٥ = ١,٧٤٠ سنتمپتراً (١٫٧٥ سم تقريبا)

طول نصف قطر الدائرة الثالثة = ١,٢٠ × ١,٥ = ١,٨٠ سنتميتراً

وأنصاف الأقطار الأخيرة التى رسمت على أساسها الدوائر فى الشكل رقم (٣-١٨) وسواء استخدمنا أى من الطريقتين السابقتين لمعرفة أطوال نصف قطر الدوائر فيجب أن لا نكتب عليها أية أعداد للكميات الحقيقية التى تمثلها الدوائر بنفس طريقة رسم الدوائر السابق شرحها ومنه يمكن أن نقيس قطر أى دائرة مرسومة، وليس من الضرورى أن يمثل هذا المقياس الدوائر المرسومة نفسها وإنما يمثل مقياساً لدوائر كمياتها ذات أرقام صحيحة بحيث تكون قريبة من الكميات الحقيقية التى تم تمثيلها بيانياً. وفى المثال السابق فان هذه الكميات تكون ١٠٠، ١٥٠، ٢٠٠ (شكل رقم: ٣-١٨).

ويمكن تقسيم الدائرة إلى أقسام داخلية للمقارنة بين أجزاء الظاهرة. وفى هذه الحالة تحول الأرقام المطلقة إلى أرقام نسبية عن طريق قسمة رقم كل جزء على المجموع الكلى للظاهرة وضربه فى ١٠٠، ثم ضرب الناتج فى ٣,٦ فنحصل على

الزاوية المركزية التي تمثل هذا الجزء. وفي المثال السابق تقسم الدائرة الأولى (٦٢/٦١) إلى ثلاثة أقسام بنسبة أنتاج كل من البحار والبحيرات والنيل كمايلي:



شكل رقم (٣-١٨)
إنتاج مناطق الصيد في مصر ٦١ - ١٩٦٤ (طريقة الدوائر المقسمة)

نسبة إنتاج البحـــار = $\frac{٦٥}{١٢٧}$ × ١٠٠ = ٥١,٢٪

نسبة إنتاج البحيرات = $\frac{٦٥}{١٢٧}$ × ١٠٠ = ٣٧٪

نسبة إنتاج النـــيل = $\frac{١٥}{١٢٧}$ × ١٠٠ = ١١,٨

وتكون الزواية المركزية لكل منها هى:

إنتاج البحـــــار = ٥١,٢ × ٣,٦ = ١٨٤°

إنتاج البحـــيرات = ٣٧ × ٣,٦ = ٢٢٣°

إنتاج النـــــيل = ١١,٨ × ٣,٦ = ٤٢°

وبالمثل يمكن حساب النسبة المئوية والزواية المركزية للإنتاج المصايد فى السنتين الأخيرتين كما فى الجدول التالى:

جدول رقم (٣-٣)

تحديد الزواية المركزية فى الدائرة الممثلة لإنتاج المصايد المصرية

السنوات / الإنتاج	٦٢/٦١		٦٣/٦٢		٦٤/٦٣	
	٪	الزاوية المركزية	٪	الزاوية المركزية	٪	الزاوية المركزية
إنتاج البحار	٥١,٢	١٨٤	٤٨	١٧٢,٨	٤٨,٦	١٧٥
إنتاج البحيرات	٣٧,٠	١٣٣	٤٠	١٤٤,٠	٣٨,٩	١٤٠
إنتاج النيل	١١,٨	٤٢	١٢	٤٣,٢	١٢,٥	٤٥
المجموع	١٠٠	٣٦٠	١٠٠	٣٦٠,٠	١٠٠	٣٦٠

ولتسهيل المقارنة بين أجزاء الظاهرة خلال فترة السنوات الثلاث يجب أن نأخذ الضلع الرأسى المكون للربع الأول من الدائرة (الخط الواصل بين مركز الدائرة وبداية تقسيم الدائرة أو صفر التدريج) ونعتبره كخط أساس يبدأ منه قياس الزوايا المركزية بعد تجميعها تصاعدياً. (شكل رقم : ٣-١٨).

Converted by Tiff Combine - (no stamps are applied by registered version)

٥- الرسوم البيانية المثلثية:

تستخدم الرسوم البيانية المثلثية فى تمثيل البيانات النسبية الخاصة بثلاث ظاهرات مختلفة، أو فى تمثل البيانات الأساسية الخاصة بثلاثة عناصر لظاهرة واحدة (مثل بيانات العمالة فى المصانع، أنواع الحيوانات، أنواع المحاصيل، أنواع النباتات، عناصر التربة) وذلك لمعرفة النسبة الغالبة بين الظاهرات أو الصفة السائدة بين عناصر الظاهرة بوجه عام.

وتقوم فكرة هذه الرسوم على أساس رسم مثلث متساوى الأضلاع، يقسم كل ضلع منه إلى عشرة أقسام متساوية تستخدم كمقياس نسبى يبدأ من الصفر حتى ١٠٠ ٪ ويكون التقسيم في إتجاه عقرب الساعة. أو بمعنى آخر أن يكون رقم ١٠٠ على أحد الأضلاع هو رقم الصفر للضلع المجاور والعكس مع رقم الصفر فيكون رقم ١٠٠ ٪ للضلع المجاور .. وهكذا. ويعد ذلك فصل بين كل رقم على أحد الأضلاع والرقم على الضلع المجاور ليكون مجموع هذين الرقمين ١٠٠ ٪ وبذلك نحصل على مجموعة من المثلثات الداخلية كل منها يشابه المقياس الكبير، ومنها نجد أن مجموع النسب لثلاثة عناصر إذ أضيفت لبعضها لحصلنا على الرقم ١٠٠ ٪ الذى يمكن تمثيله على الرسم البيانى المثلثى بنقطة واحدة فقط. وفى عملية توقيع مكان هذه النقطة نبحث أولا عن القيمة المراد تمثيلها على أحد الأضلاع الذى يمثل أحدى العناصر، ونمد منها خطاً يلتقى مع الخط الذى يمد من مكان القيمة الثانية على أحد الضلعين الآخرين. وسيلقى حتماً كل من الخطين مع الخط الواصل عن مكان القيمة الثالثة على الضلع الثالث. ومكان تلاقى الخطوط الثلاثة هو موقع النقطة التى ستجمع قيم الثلاث ظاهرات أو الثلاثة عناصر فى موضع واحد على الرسم.

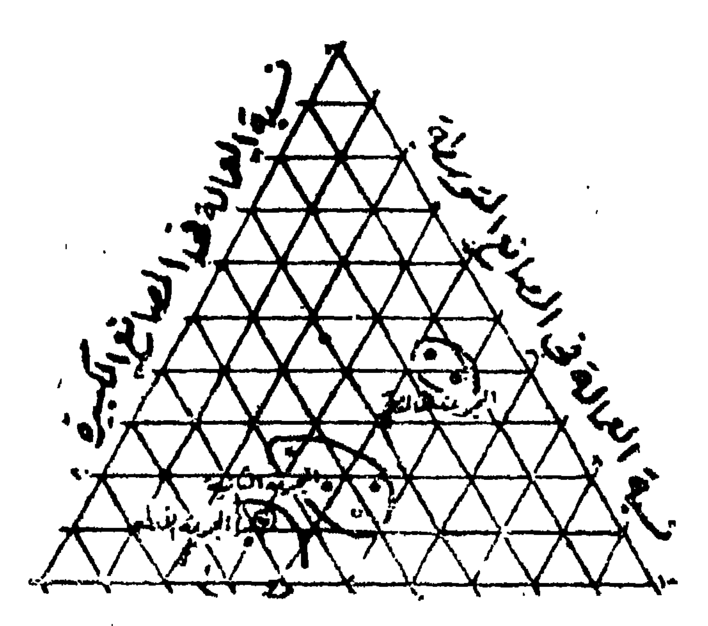
ومن الأمثلة الشائعة لاستخدام هذه الرسوم مايختص بتحليل عينات التربة فمثلاً إذا كانت لدينا مجموعة من عينات التربة وكان تحليلها على أساس النسب المئوية للعناصر الثلاثة الرئيسية التى تتألف منها (الرمل، الغرين، الصلصال) وكان المطلوب تمثيلها بيانيا لمعرفة الصفة الغالبة للتربة بوجع عام، كان من الممكن

عندئذ استخدام هذا النوع من الرسوم البيانية. وبالمثل يمكن استخدامه لبيان الحالة العامة لثلاث من أنواع الصناعات فى مجموعة من المدن.

ولهذه الرسوم البيانية أهمية خاصة اذ أنه يمكن أن تتخذ كأساس لوضع تصنيف أو استخراج أنماط لعديد من الظاهرات عن طريق تحديد بعض المساحات على الرسم والتى منها يمكن أن نعرف موقع القيمة الثلاثية. فإذا كان هذا الموقع بالقرب من أحد أركان (نقط) المثلث يعنى أن قيمة أحد العناصر (احدى الظاهرات) لابد أن تكون كبيرة جداً، بينما وقوع القيمة الثلاثية بالقرب من جوانب المثلث يشير إلى أن قيمة أحد العناصر (احدى الظاهرات) لابد أن تكون صغيرة جداً.

وكمثال تطبيقى يمكن سرده لبيان استخدام هذا النوع من الرسوم البيانية استخدمت بيانات العمالة فى ١٢ مصنعاً رئيسياً بالسويد. وقسمت هذه المصانع حسب حجم العمالة المدربة بها (١٠٠ - ٥٠٠ عامل) إلى ثلاث فئات هى: مصانع صغيرة، مصانع متوسط الحجم، مصانع كبيرة. والشكل رقم (٣-١٩) يوضح تمثيل بيانات الفئات الثلاث من المصانع، ومنه يمكن أن نرى أن الصناعات السويدية يمكن أن تصنف إلى ثلاث مجموعات (بدون الصناعات الهندسية والحديدية التى تقف كمجموعة بمفردها) : المجموعة الأولى تشتمل على التحجير، الطباعة، الأعمال الخشبية، الصناعات الغذائية والمشروبات. وصناعات هذه المجموعة يقوم بانتاجها نسبة كبيرة من العمالة المدربة فى المصانع الصغيرة لا تقل عن ٥٥٪ من جملة العمالة فى الأنواع الثلاثة من المصانع، بينما تقل نسبة العمالة المدربة فى المصانع الكبيرة فى مجال إنتاج صناعات هذه المجموعة. والمجموعة الثانية تشمل الصناعات الجلدية، الغاز، المياه، والكهرباء، الصناعات الكيميائية، والنسيج. وهى صناعات لا يتعادل التدريب فى إنتاجها بين الأنواع الثلاثة من المصانع حيث نجد أن هناك نسبة عمالة مدربة صغيرة للمصانع الكبيرة بينما تتعادل تقريباً نسبة العمال التى تدربها المانع المتوسطة الحجم والمصانع الصغيرة التى تهتم بأنشطة المجموعة الصناعية الثانية. أما المجموعة الثالثة فتمثل الصناعات التعدينية وتصنيع الورق، وهذه تتساوى فيها تقريباً نسبة العمالة المدربة فى المصانع الكبيرة والمتوسطة

الحجم، بينما تقل نسبة العمالة المدربة كثيراً لإنتاج صناعات هذه المجموعة فى المصانع الصغيرة الحجم.

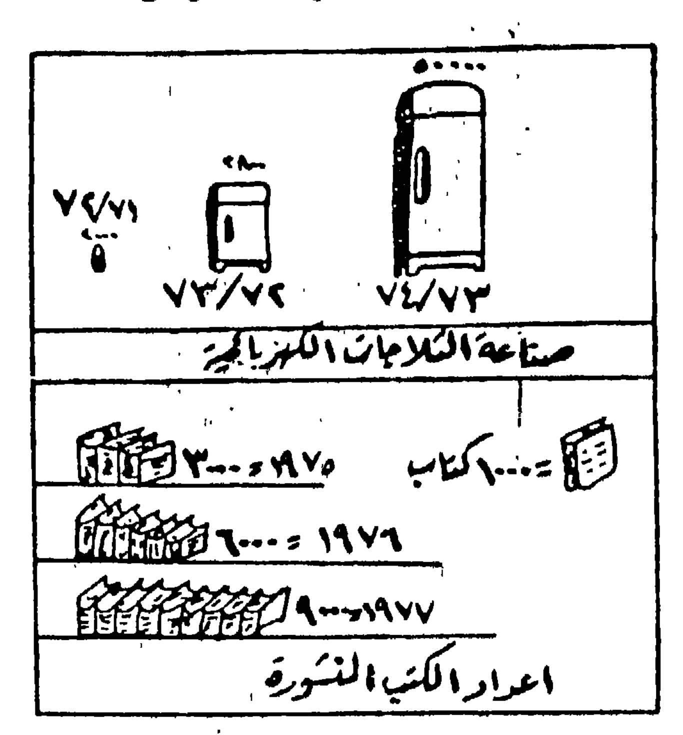


شكل رقم (٣-١٩)
تحديد فئات المصانع ونسبة العمالة بها
«طريقة الرسوم البيانية المثلثية»

ثالثاً: الطرق البيانية بالرسوم التصويرية Pictorial graphs

تعد طريقة الرسوم البيانية التصويرية من أحسن وسائل الإيضاح وأكثرها جاذبية فى التعبير عن تغير وتطور كميات الظواهر. وتقوم فكرة هذه الطريقة على أساس إعطاء وحدات قياس (مفردات) الظاهره أشكالاً تصويرية (شكل رقم ٣-٢٠). فمثلاً عدد السكان يمكن أن نمثله برسم عدد من الاشخاص يمثل الشخص الواحد عدد مليون نسمة، أو عدد الثلاجات لأحد مصانع الثلاجات يمكن تمثيله برسم عدد من الثلاجات كل واحدة منها تمثل عشرة أو مائة ثلاجة. وكذلك عدد الكتب التى تصدرها احدى دور النشر تصور بكتاب لكل

عـدد مـعين من هذه الكتب (شكل رقم ٣-٢١)، كـمـا أن عـدد قـراء إحـدى الصحف اليومية يصور على أساس قارئ لكل عدد معين من الأعداد الصادرة.



شكل رقم (٣-٢١)
نموذج للرسوم البيانية التصويرية

وفى كل الحالات السابقة يجب مراعاة الدقة فى الرسم حتى يمكن الحصول على عدد من الرسوم التوضيحية أو الرموز تشابه تماماً الظاهرة المراد تمثيلها وعند تحديد عدد الوحدات من الظاهرة والتى يمثلها الشكل المختار فانه يجب تحديد هذا العدد فى ضوء أكبر وأصغر رقم فى البيانات وكذلك فى ضوء مساحة اللوحة المخصصة للرسم. ويمكن كذلك تكبير أو تصغير الوحدة التى تمثل الظاهرة ويراعى أن يكون هذا التكبير والتصغير على أساس مقياس الرسم لعدد المرات التى تساويها الوحدة الصغيرة. أما كسور القيم فيمكن تمثيلها برموز أو أشكال غير كاملة.



شكل رقم (٣-٢٢)

أعداد قراء الصحف اليومية «طريقة الرسوم التصويرية»

ويعاب على هذه الطريقة رغم أنها تشد أنظار الشخص العادى فى التعرف على طبيعة الظاهرة من حيث تطورها وتغيرها، إلا أنها لا تعطى فكرة دقيقة عن قيم الظاهرة حيث أننا قد نضطر أحياناً إلى تقريب القيم إلى أقرب عشرة أو مائة أو ألف حتى يمكن التخلص من الكسور الصغيرة والتى لا يمكن أن نعطى لها شكلاً أو رمزاً. فمثلاً القيمة ٦٠٠ كتاب من الصعب تمثيلها تصويرياً برمز أو شكل متكامل إذا كان الأخير يمثل ١٠٠٠ كتاب مثلاً. كذلك تلائم هذه الطريقة تمثيل البيانات التى تتميز بتباين مفردات قيمها. وأيضاً لا تعطى هذه الطريقة فكرة حقيقية واضحة عن حقيقة المقارنة بين هذه الوحدات.

العرض البيانى للبيانات المبوبة:

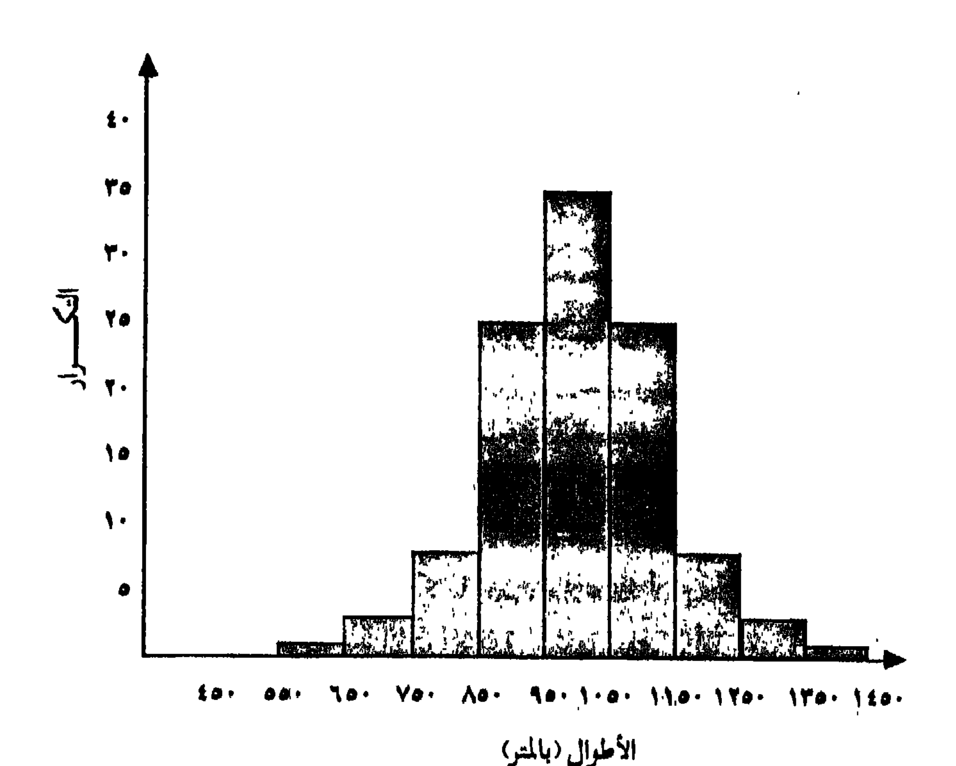
يقصد بالعرض البيانى للبيانات المبوبة هو التمثيل التصويرى لبيانات جداول التوزيعات التكرارية. ويأخذ هذا العرض البيانى أشكالاً مختلفة الهدف منها إعطاء صورة إجمالية عن خصائص البيانات وتوضيح الصفات المميزة للتوزيعات التكرارية.

وعموماً فإن بيانـات جداول التوزيعات التكرارية تمثل بيانيا بعدة طرق مختلفة أهمـهـا المدرج التكرارى Histogram، المضلـع التكـرارى Polygon والمنحنى التكرارى Frequency Curve..

المدرج التكرارى Frequency Histogram

يتكون المدرج التكرارى من مجموعة من المستطيلات المتلاصقة التى تكون قواعدها على المحور الأفقى (محور الستينات) متساوية وطول كل قاعدة منها يساوى طول فترة الفئة، كما يتناسب ارتفاع كل مستطيل مع التكرار المناظر لكل فئة على المحور الرأسى. ويراعى عند رسم المدرج التكرارى إختيار مقياس رسم ملائم لكل محور، بحيث تمثل جميع الفئات على المحور الأفقى كما يسمح المحور الرأسى بتمثيل جميع التكرارات. كما يجب ملاحظة أن يبدأ المحور الرأسى دائماً بالصفر حتى يمكن مقارنة التكرارات المختلفة بعضها ببعض. أما المحور الأفقى فليس من الضرورى أن يبدأ بالصفر، إذ أن ذلك سوف يحتاج منا إلى محور طويل جداً لتمثيل الفئات.

وإذا أردنا تمثيل التوزيع التكرارى لأطوال مائة رافد نهرى (جدول رقم ٢-١١) برسم مدرج تكرارى له، نلاحظ أن التوزيع منتظم وينقسم إلى ٩ فئات متساوية، ولذلك نقوم بتقسيم المحور الأفقى إلى ٩ أقسام متساوية (ويستحسن أن نزيد عليها قسماً أى تصبح ١٠ أقسام). ويفضل أن نبدأ تقسيم المحور الأفقى بفئة أصغر من أقل فئة بالجدول أى نبدأ بتقسيم المحور الأفقى بالعدد ٤٥٠. ولما كان أكبر التكرارات هو ٣٣ فإننا نقوم بتقسيم المحور الرأسى إلى أقسام متساوية بمقياس رسم مناسب بحيث يسمح بظهور أكبر التكرارات بالجدول بحيث يبدأ التقسيم على المحور بالصفر. ثم نرسم على كل فئة من الفئات مستطيلاً تكون قاعدته مساوية لطول فترة الفئة وارتفاعه يساوى تكرار الفئة المقام فوقها هذا المستطيل. وبذلك نحصل على المدرج التكرارى الموضح فى شكل رقم (٣-٢٣)، والذى فيه مساحة المستطيلات تتناسب مع التكرارات (لأن أطوال قواعد المستطيلات جميعها متساوية ولأن مساحة كل مستطيل تساوى طول القاعة فى الارتفاع).



شكل رقم (٣-٢٣)
المدرج التكرارى لأطوال ١٠٠ رافد نهرى (بالمتر)

واذا كانت الفئات كلها لها نفس الطول، فإنه من المعتاد أن ترسم الارتفاعات بأطوال مساوية لتكرارات الفئات. أما إذا كانت الفئات غير متساوية الطول فإن هذه الأطوال يجب أن تعدل قبل رسم المدرج التكرارى. وذلك بقسمة تكرار كل فئة على طول هذه الفئة وبذلك نحصل على التكرار المعدل، ثم ننشئ المدرج التكرارى برسم عدد من المستطيلات قاعدتها تمثل أطوال الفئات «غير المتساوية» على المحور الأفقى وارتفاعاتها تمثل التكرارات المعدلة. وفى هذه الحالة تتناسب مساحة كل مستطيل مع التكرارات المعدلة المناظرة لكل فئة.

ومما تجدر الإشارة إليه أن المدرج التكرارى يصلح لتمثيل بيانات المتغيرات

المستمرة «المتصلة» ولايصلح لتمثيل بيانات المتغيرات غير المستمرة «الوثابة»، إذ يجب أن تكون حدود فئات التوزيع المرسوم له الهيستو جرام معروفة تماماً. كما أن المدرج التكرارى يعبر عن اتجاهات لتوزيع التكرارى عن طريق إبراز الصفات العامة للتوزيع.

المضلع التكرارى Frequency Polygon

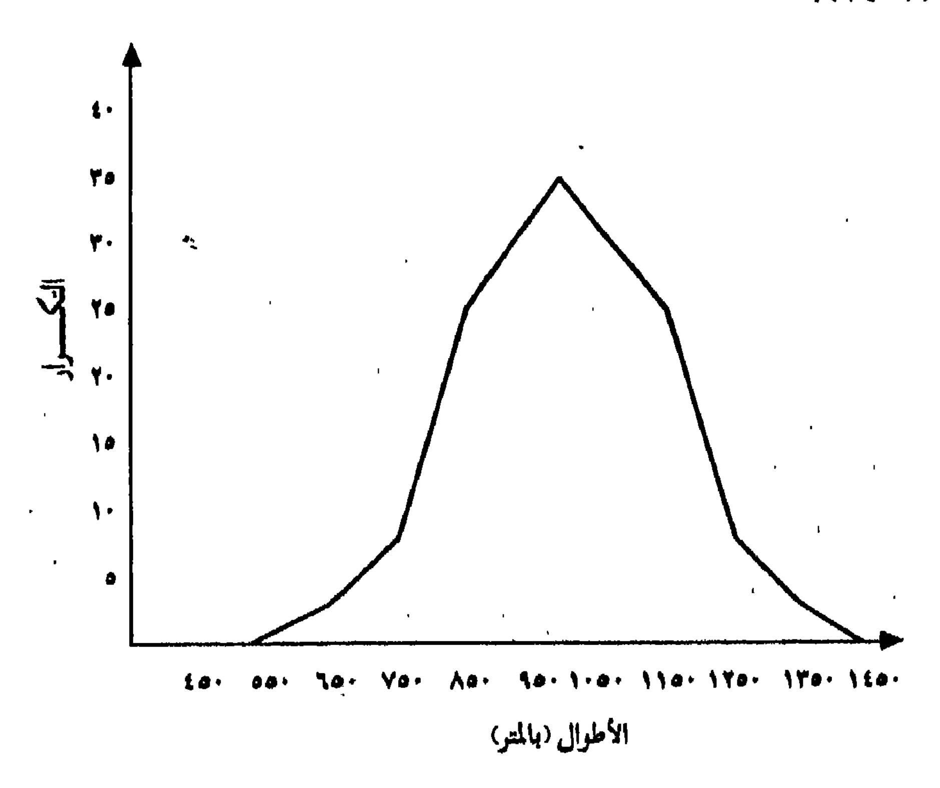
يمكن اعتبار إنشاء المدرج التكرارى بالطريقة التى ذكرناها سابقاً خطوة من خطوات رسم المضلع التكرارى للتوزيع التكرارى. وتعتمد فكرة إنشاء المضلع التكرارى على فكرة التمثيل البيانى من خلال الخط البيانى حيث تبين فقط تمثيل تكرار الفئة المقابلة لمركز الفئة. ويرسم المضلع التكرارى بإيصال نقط تنصيف المستطيلات المكونة للمدرج التكرارى بمجموعة من الأضلاع يقع كل ضلع منها بين مركز فئة ما ومركز الفئة التالية مباشرة. وحتى يمكن قفل شكل المضلع التكرارى وتحديد مساحته فإننا نفترض وجود فئتين إحداهما قبل الفئة الأولى والأخرى بعد الفئة الأخيرة والتكرار المناظر لكل منهما يساوى صفراً، ثم يتم توصيلهما بأضلاع مع مركزى الفئة الأولى والأخيرة فتحصل بذلك على شكل كامل للمضلع التكرارى كما فى شكل رقم (٣-٢٤).

وفكرة إستخدام مركز الفئة أو منتصفها فى رسم المضلع التكرارى يعتمد على افتراض تركز كل القيم عند متوسطها الحسابى حيث أن مركز الفئة يساوى:

$$\frac{\text{الحد الأدنى للفئة} + \text{الحد الأعلى للفئة}}{٢}$$

ويتميز المضلع التكرارى بأن المساحة المحصورة تحت المضلع هى نفسها المساحة المحصورة تحت المدرج التكرارى، ولكنه يكون أكثر دقة من المدرج التكرارى من حيث اعطائه صورة أكثر واقعية لاتجاهات وخصائص التوزيع. كما يعد المضلع التكرارى من أنسب الطرق البيانية لتمثيل أكثر من توزيع واحد من التوزيعات التكرارية مما يسهل إجراء المقارنة بينها. ويجب ملاحظة أنه لا يمكن إيجاد تكرار أية قيمة تقع بين مركزى فئتين من المضلع التكرارى وذلك لأن المضلع التكرارى

يمثل تكرار الفـئـات لا تكرار القـيم الفـردية كـمـا هو واضح فى الشكل رقـم (٣-٢٤).



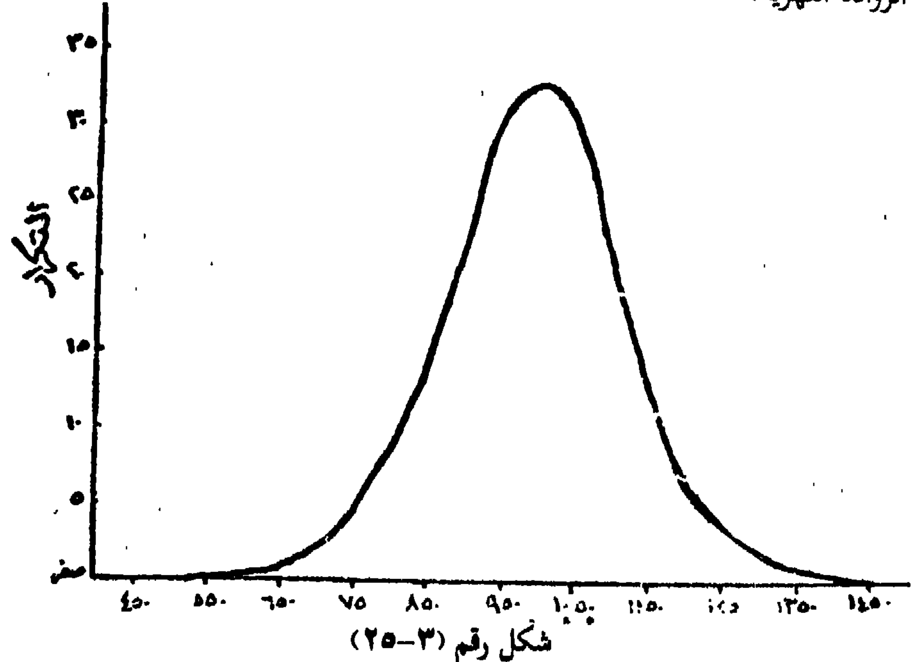
شكل رقم (٢-٢٤)
المضلع التكرارى لأطوال ١٠٠ رافد نهرى «بالمتر»

المنحنى التكرارى: Frequency Curve

يمكن تمثيل التـوزيعـات التكرارية بطريقـة أخرى، غيـر الطريقين السابقين، تظهـر فى شكل هندسى واضح وذلك برسم المنحنى التكرارى للتـوزيع والذى نحـصل عليـه من خـلال رسم المضلع التكرارى أولا ثم تمـهـيـد خطوط المضلـع المنكسرة. وتقـوم فكرة المنحنى التكرارى على أساس اعتبار أن البيانات تمثل عينة مسحوية من مجتمع أكبر. وبما أن هناك عدداً كبيراً من المفردات فى المجتمع فإنه من الممكن من الناحية النظرية (للبيانات المستمرة) إختيار فترة الفئة صغيرة جداً

ويظل لدينا عدد ملموس من المفردات (تكرارات) فى داخل كل فئة. وبهذا فإنه من المتوقع أن يتكون المضلع التكرارى للمجتمعات الكبيرة من عدد كبير من الخطوط الصغيرة المنكسرة والتى يمكن تقريبها بمنحنى. وتزيد درجة الدقة فى التقريب بزيادة حجم العينة. ولهذا السبب فإن المنحنى التكرارى يسمى أحياناً المضلع التكرارى الممهد.

ولرسم المنحنى التكرارى نعين مراكز الفئات على حسب التكرارات المناظره لها، ونوقع نقط المضلع التكرارى ونمهد الخطوط المنكسرة بين هذه النقط باليد بحيث يمر المنحنى بأغلبية رءوس المضلع التكرارى. (يوجد أيضاً طرقا رياضية لتكوين المنحنى التكرارى تجعل المساحة تحت المنحنى مساوية للمساحة تحت المضلع التكرارى). والشكل رقم (٣-٢٥) يبين المنحنى التكرارى لتوزيع أطوال الروافد النهرية.



شكل رقم (٣-٢٥)

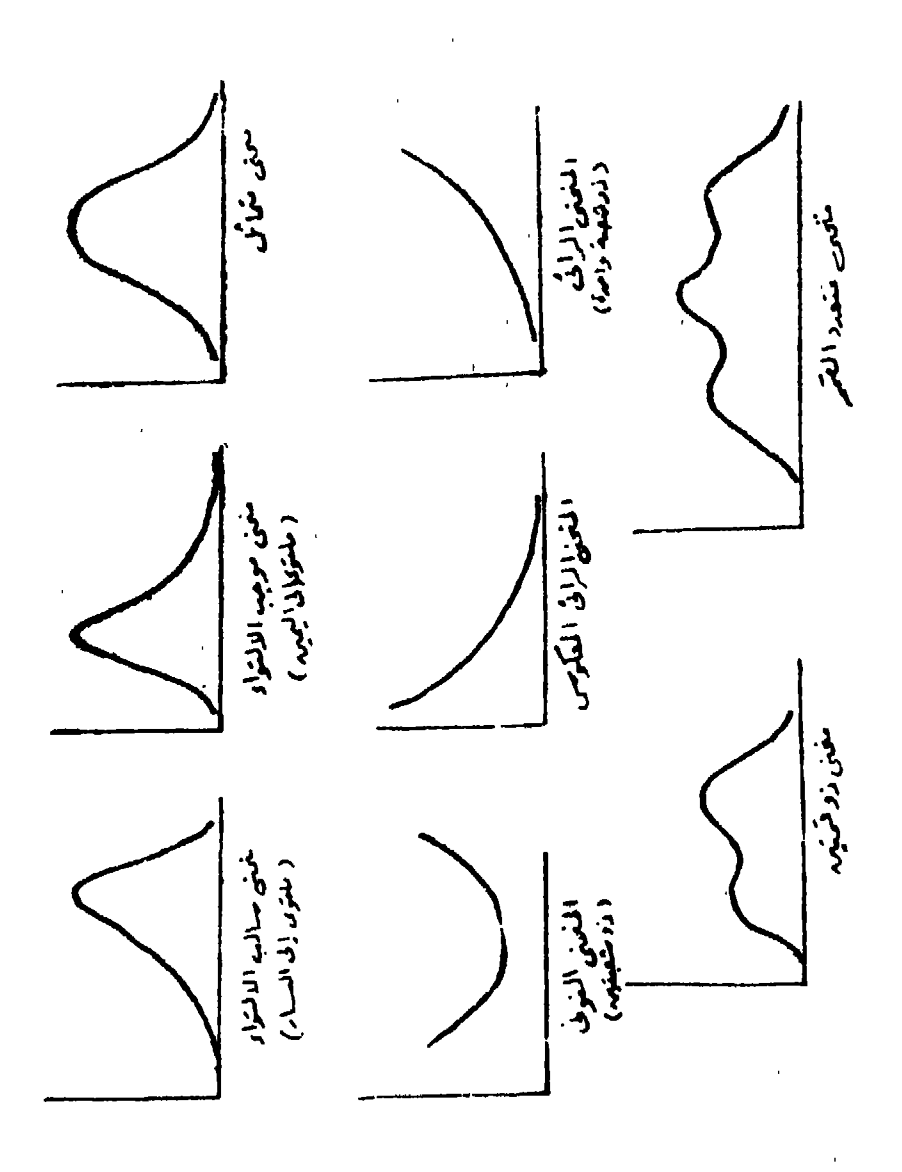
المنحنى التكرار لأطول ١٠٠ رافد نهرى (بالمتر)

ونظراً لأن المنحنى التكرارى يمكن رسمه من نقط المضلع التكرارى، لذا فإن شكل المنحنى يتوقف على توزيع البيانات ومن المنطقى أن نتوقع وجود عدد كبير من الأشكال المختلفة للمنحنيات التكرارية، إلا أنه يمكن حصر أشكال المنحنيات التى تقابلنا عادة فى التحليل الاحصائى للبيانات، والتى يمكن تقسيمها إلى نوعين رئيسيين:

النوع الأول: هو المنحنيات المتماثلة Symmetrical Curves. وهى المنحنيات التى تتماثل حول خط رأسى يقسم المنحنى إلى قسمين متطابقين أو بمعنى آخر هى المنحنيات التى تتماثل فيها التكرارات على جانبى أكبر تكرار. أما النوع الثانى: فهو المنحنيات غير المتماثلة Asymmeterical Curvese وهى المنحنيات التى تظهر فيها صفة التماثل كأن تبدو فيها التكرارات متناقصة فى طرف عنه فى الطرف الآخر. ويأخذ كل نوع من هذين النوعين من المنحنيات أشكالاً مميزة كما هو كما هو موضح بالشكل رقم (٣-٢٦) ونعرضها فيمايلى:

١- المنحنى التكرارى المتماثل Normal Curve وهو أكثر المنحنيات شيوعاً ويأخذ هذا المنحنى الشكل الناقوسى Bell- shaped Curve الذى يتميز بأن المفردات المتساوية البعد عن مركز النهاية العظمى لها نفس التكرارات. ومن الأمثلة الهامة له المنحنى المعتدل وهو المنحنى الذى نتوقع الحصول عليه من دراسة كثير من الظاهرات التى تتغير تبعاً لأسباب طبيعية مثل الطول والوزن..... الخ. ونظراً للأهمية الكبيرة للمنحنى المعتدل فى الدراسات الإحصائية لذا سندرسه بالتفصيل فيما بعد.

٢- المنحنيات التكرارية الملتوية Skewed Culves وهى المنحنيات التى تتميز بأن أحد طرفيها يمتد أكثر من الآخر على جانبى مركز النهاية العظمى. فإذا كان الطرف الأيمن للمنحنى أطول من الطرف الأيسر فإن المنحنى يكون فى هذه الحالة ملتوياً إلى اليمين أو ملتوياً إلتواء موجباً Positively Skewed- Curve بينما لو كان العكس صحيحاً فأن المنحنى يكون ملتوياً إلى اليسار أو ملتوياً إلتواء سالباً Negativcly Skewcd- Curve.



شكل رقم (٣-٢٦) أشكال المنحنيات التكرارية

٣- المنحنيات ذات الشكل الرائى أو الشكل الرائى المعكوس. وفيها تقع نقطة النهاية العظمى عند أحد طرفى المنحنى. ويطلق على هذين النوعين من المنحنيات اسم المنحنيات الأسية Exponcntial الموجبة والسالبة على الترتيب. ومن أمثلة هذه المنحنيات منحينات توزيع الثروة أو الملكية فى المجتمع.

٤- المنحنى النونى، وهو المنحنى الذى يأخذ شكل حرف « u » U-shaped Curve. ويتميز هذا المنحنى بأن له نهاية عظمى (التكرارات الكبرى) عند كل من طرفيه بينما تكون النهاية الصغرى أو التكرارات الأقل فى اتجاه مركز المنحنى. ويعتبر توزيع الوفيات للسكان حسب السن من أوضح أمثلة هذا الشكل من المنحنيات.

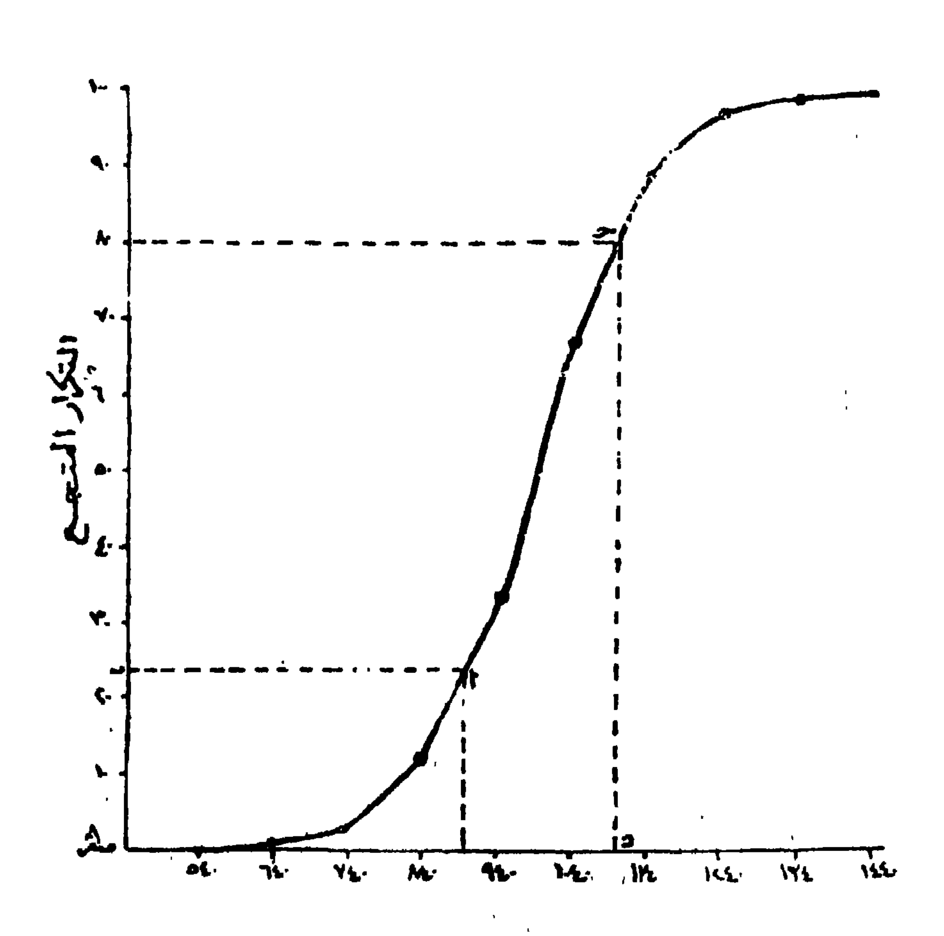
٥- المنحنى ذو القمتين Bimodal Curve، وهو المنحنى الذى يتميز بأن له نهايتان عظيمتان، والسبب فى ذلك يرجع عادة إلى عدم تجانس العينة موضع الدراسة فقد تحتوى على مجموعتين مختلفتين ومتداخلتين من المفردات.

٦- المنحنى متعدد القمم Multi- mode، وهو المنحنى الذى له أكثر من نهايتين عظيمتين ويدل على عدم تجانس المفردات فى العينة بمعنى أن تكون العينة مكونة من خليط من المجموعات.

المنحنى التكرارى المتجمع: Cumulative Frequency Curve (Ogive)

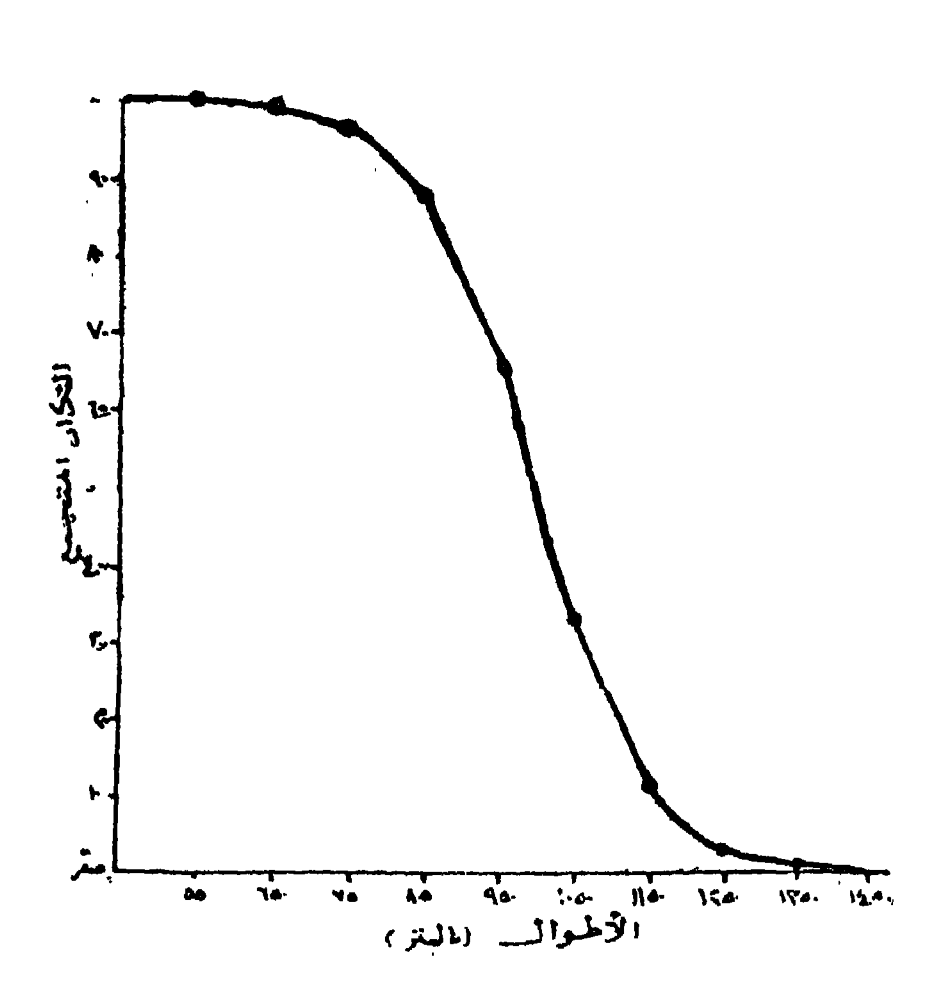
إلى جانب الأنواع السابقة من المنحنيات التكرارية هناك منحنيات بيانية تمثل التوزيعات التكرارية المتجمعة. وقد سبق أن أوضحنا كيفية عمل جداول التكرارات المتجمعة الصاعدة والهابطة وكذلك التكرارات المتجمعة النسبية. وتعرض جداول التوزيع التكرارى المتجمع بيانياً على هيئة مضلع تكرارى متجمع أو منحنى تكرارى متجمع Ogive وهو إما من النوع الصاعد (أقل من) أو من النوه الهابط (أكبر من). ففى الحالة الأولى تمثل الحدود العليا للفئات على المحور الأفقى والتكرارات المتجمعة على المحور الرأسى، وتوقع القيم حسب إحداثيها السينى والصادى بنقط فصل بينها بمنحنى ممهد فنحصل على المنحنى المتجمع الصاعد (شكل رقم ٣-٢٧) لأن التكرارات المتجمعة تكون فى تزايد. أما فى الحالة الثانية فتمثل

الحدود الدنيا للفئات الأصلية على المحور الأفقى والتكرارات المتجمعة الهابطة على المحور الرأسى، ثم نصل بين النقط التى تمثل القيم فى الجدول بمنحنى ممهد فنحصل على المنحنى المتجمع الهابط، كما فى الشكل رقم (٣– ٢٨).



شكل رقم (٣–٢٧)

المنحنى المتجمع الصاعد لعدد ١٠٠ رافد نهرى بالمتر



شكل رقم (٣-٢٨)

المنحنى المتجمع الهابط لعدد ١٠٠ رافد نهرى (بالمتر)

وتتميز المنحنيات التكرارية المتجمعة (الصاعدة والهابطة) بأنه يمكن إستخدامها لإيجاد تكرار أى قيمة وسيطية تقع مابين مركزى فئتين، وكذلك إيجاد الحد الأعلى لأى فئة يقع بداخلها تكرار معين. فمثلاً إذا أردنا معرفة عدد الروافد التى تقل أطوالها عن ٩٠٠ متر فإننا نقيم عموداً على المحور الأفقى عند القيمة ٩٠٠ حتى يقابل المنحنى المتجمع الصاعد فى تقطة (أ) ثم نرسم منها خطاً أفقياً موازياً للمحور الأفقى حتى يلاقى المحور الرأسى فى نقطة (ب) تحدد لنا عدد الروافد المطلوبة وهو ٢٤ (شكل رقم ٣-٢٧). ويحدث العكس إذا أردنا معرفة الحد الأعلى لأطوال ٨٠ رافداً فإننا نرسم خطاً أفقياً موازياً للمحور الأفقى عند ٨٠ حتى

يقابل المنحنى فى نقطة (حـ) ثم نسقط منها عموداً على المحور الأفقى فيقطعه فى نقطة (د) تمثل الحد الأعلى لأطوالها وهو ١١٠٠ مترا من الرسم. أى أن هناك ٨٠ رافداً نهرياً أطوالها أقل من ١١٠٠ متراً.

وفى حالة إذا استخدمنا التكرارات المتجمعة النسبية لرسم منحنى تكرارى (صاعد أو هابط) بنفس الطريقة المتبعة فى رسم المنحنيات المتجمعة، فإن النتيجة تسمى بالمنحنى التكرارى المتجمع النسبى Camulative precentage frequency curve

منحنى لورنز Lorenz Curve

يعد منحنى لورنز كأحد مؤشرات التفاوت Index of difference أو التركيز من أهم تطبيقات منحنى التوزيع المتجمع الصاعد التسبى. وحيث أن التركيز هنا يعنى سوء أو عدم عدالة التوزيع Inequality ، فإن الغرض من إستخدام منحنى لورنز هو بيان درجة التفاوت فى توزيع الظاهرات ومقارنة عدالة توزيع لهذه الظاهرات من مكان لآخر أو من فترة زمنية إلى أخرى لنفس المكان. ومن أوضح نماذج البيانات الجغرافية التى يستخدم فى تحليلها منحنى لورنز بيانات توزيع ملكية الأراضى الزراعية وتوزيع العمالة على الصناعات المختلفة.

ويعتمد أسلوب منحنى لورنز فى التحليل على وجود توزيع نظرى (أمثل) Hypothetetical or even distribution للظاهرة موضع الدراسة، ويتمثل هذا التوزيع بخط مستقيم تعطى أى نقطة عليه النسب المتساوية فى التوزيع. فمثلاً إذا كنا بصدد قياس علاقة مساحة الأراضى الزراعية بالملاك فإنه من المفروض (نظرياً) أن نسبة ١٠ ٪ من الملاك يمتلكون ١٠ ٪ م الأراضى كما أن ٧٠ ٪ من الملاك يمتلكون ٧٠ ٪ من الأراضى.. وهكذا. كما يعتمد أسلوب منحنى لورنز أيضاً على وجود منحنى للتوزيع الفعلى للظاهرة. ولما كان التوزيع الفعلى لأى ظاهرة يختلف حتماً عن التوزيع النظرى (المثالى) لها، فإن العلاقة بين منحنى التوزيع الفعلى وخط التوزيع المتساوى تحدد درجة عدالة أو سوء توزيع هذه الظاهرة. ويستدل على ذلك من المساحة المحصورة بين المنحنى وخط التوزيع المتساوى

(مساحة التركيز أو سوء التوزيع)، فكلما قرب المنحنى من هذا الخط كلما صغرت المساحة بينهما وذل ذلك على قرب التوزيع من التوزيع المثالى أو اتخذ دليلاً على عدالة التوزيع، أما إذا بعد منحنى التوزيع الفعلى عن خط التوزيع المتساوى فإن المساحة المحصورة بينهما تزداد ويدل ذلك على بعد التوزيع عن التوزيع المثالى أو يتخذ دليلاً على سوء توزيع الظاهرة.

ولرسم منحنى لورنز لتوزيع الملكية الزراعية مثلا (جدول رقم ١٣) نتبع الخطوات التالية:

١- تحول التكرارات الأصلية التى تمثل عدد الملاك والمساحة المملوكة إلى تكرارات متجمعة صاعدة.

٢- تحول التكرارات المتجمعة المطلقة إلى تكرارات متجمعة نسبية صاعدة وبالتالى نحصل على التوزيع الفعلى الذى يتم تعيينه على الرسم البيانى.

٣- نرسم مربعاً يمثل أحد محوريه الأفقين التكرارات المتجمعة النسبية لعدد الملاك ويمثل أحد محوريه الرأسيين التكرارات المتجمعة النسبية لجملة المساحة، بحيث يبدأ مقياس كل محور بالصفر وينتهى برقم ١٠٠ ٪. توقع النقط الممثلة للنسب على الرسم البيانى ونصل بينها بخط ممهد يطلق عليه منحنى التوزيع الفعلى (منحنى لورنز).

٤- نرسم الخط المستقيم الذى يصل بين نقطتى الصفر و ١٠٠ (قطر المربع) ليمثل التوزيع المثالى (النظرى) للتوزيع كما فى الشكل رقم (٣-٢٩).

جدول رقم (٣-٤)

توزيع الملكية الزراعية حسب فئات المساحة

وجملة المساحة المملوكة قبل صدور قانون الإصلاح الزراعى عام ١٩٥٢

فئات المساحة (الفدان)	عدد الملاك (أقرب ألف)	جملة المساحة المملوكة (لأقرب ألف)	التكرار المتجمع الصاعد		التكرار المتجمع النسبى الصاعد	
			٪ لعدد الملاك	جملة المساحة المملوكة	٪ لعدد الملاك	٪ الجملة المساحة المملوكة
أقل من ٥	٢٦٤٢	٢١٢٢	٢٦٤٢	٢١٢٢	٩٤,٣	٣٥,٤
أقل من ٥ —١٠	٧٩	٥٢٦	٢٧٢١	٢٦٤٨	٩٧,١	٤٤,٢
أقل من ١٠ — ٢٠	٤٧	٦٣٨	٢٧٦٨	٣٢٨٦	٩٨,٨	٥٤,٩
أقل من ٢٠ — ٥٠	٢٢	٦٥٤	٢٧٩٠	٣٩٤٠	٩٩,٦	٥٦,٨
أقل من ٥٠ — ١٠٠	٦	٤٣٠	٢٧٩٦	٤٣٧٠	٩٩,٨	٧٣,٠
أقل من ١٠٠ — ٢٠٠	٣	٤٢٧	٢٧٩٩	٤٨٠٧	٩٩,٩	٨٠,٣
٢٠٠ فأكثر	٢	١١٧٧	٢٨٠١	٥٩٨٤	١٠٠,٠	١٠٠,٠
المجموع	٢٨٠١	١١٨٤				



الباب الثانى
مقاييس الو صف



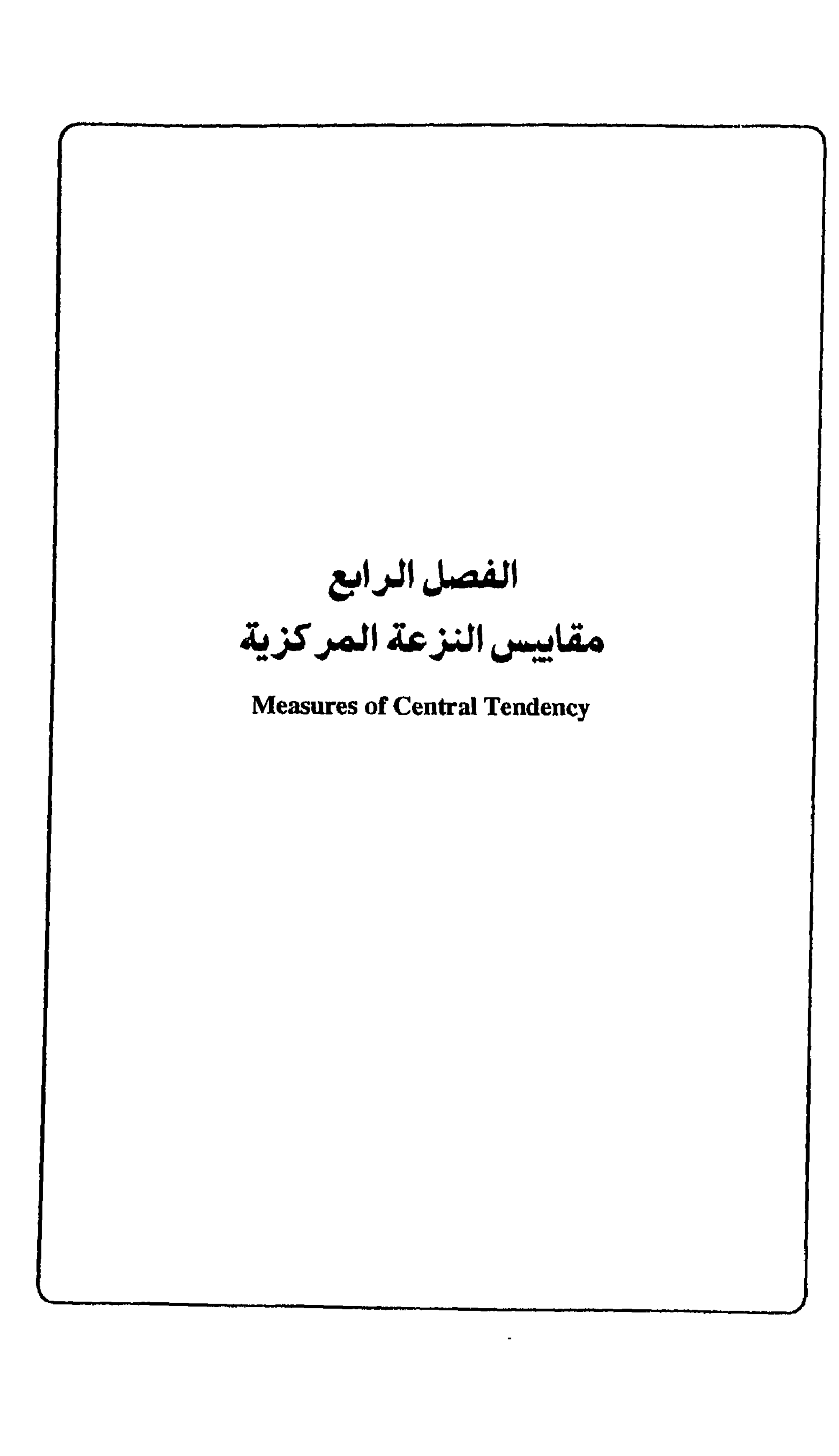
الباب الثانى
مقاييس الوصف

مقدمة:

ذكرنا فى الباب السابق أن عملية تبويب وتنظيم البيانات الإحصائية عن طريق وضعها فى جداول أو تمثلها بيانياً لاتتم بغرض عرض البيانات بصورة تلخص معالمها فحسب، بل أنها أيضاً تعد الخطوة الأولى على طريق التحليل الإحصائى لهذه البيانات. وحيث أن أسلوب الجدولة والتمثيل البيانى فى دراسة الظاهرات يعتمد على دقة الأسلوب نفسه، فإن الخطوة التالية فى التحليل هى وصف البيانات بطريقة موضوعية غير متحيزة واستخلاص النتائج منها. ونلجأ فى وصف البيانات Data Descripion إلى حساب بعض أساليب القياس الكمى والتى تعرف باسم «المقاييس الوضعية أو مقاييس الوصف الإحصائى». وهى إما مقاييس محسوبة من المجتمع وتسمى بمعالم (ثوابت) المجتمع Paramaters، أو مقاييس محسوبة من العينة وتعرف باسم إحصائيات العينة a statistic.

ونظراً لاختلاف الظواهر الجغرافية (طبيعية وبشرية) من حيث الخصائص والاتجاهات وبالتالى اختلاف نوعية البيانات وتنوع توزيعاتها التكرارية، فلقد تعددت أدوات القياس الكمى التى تهتم بتحديد هذه الخصائص والاتجاهات. فمثلاً قد تختلف التوزيعات فى القيمة المتوسطة التى تتركز حولها قيم المفردات، أو قد

أو فى اتجاه وشكل تركز قيم المتغيرات فى أحد نواحى التوزيع وتعرف المقاييس الكمية فى الحالة الأولى بمقاييس النزعة المركزية، وتعرف المقاييس فى الحالة الثانية بمقاييس التشتت (الإختلاف)، بينما تعرف المقاييس فى الحالة الثالثة بمقاييس أو مؤشرات التركز. وسنعرض فى هذا الباب أنواع مقاييس الوصف الكمية وطرق استخدامها وخصائصها ومجالات تطبيقاتها المتعددة فى دراسة البيانات الاحصائية بأنواعها المختلفة، إلى جانب عرض مفصل لمزايا استخدام هذه المقاييس ومشاكل تطبيقاتها.



الفصل الرابع

مقاييس النزعة المركزية

Measures of Central Tendency

مقاييس النزعة المركزية

المقصود بالنزعة المركزية هو نزعة المفردات للتركز حول قيمة متوسطة أو قيمة نموذجية تمثل مجموعة من البيانات. وحيث أن مثل هذه القيمة النموذجية تميل إلى الوقوع فى المركز داخل مجموعة بيانات مرتبة حسب قيمها، فقد اتخذت كأساس للوصف الإحصائى لمعالم المجموعة التى تشكلها هذه البيانات. ويمكن أن نعرف صوراً عديدة لمقاييس النزعة المركزية وإن كان من أكثرها شيوعاً: المتوسط بأنواعه (المتوسط الحسابى، المتوسط الهندسى، المتوسط التوافقى)، الوسيط والمنوال وكل مقياس من هذه المقاييس له مميزاته وعيوبه وهنا يعتمد على البيانات والهدف من استخدامه. ويطلق فى بعض الأحيان على المقاييس السابقة للنزعة المركزية اسم مقاييس المتوسطات Averages أو مقاييس الموضع Location.

أولاً: المتوسط Mean

يختلف تعريف المتوسط للبيانات تبعاً لنوع المقياس المستخدم، إلا أنه يعتبر أبسط مقاييس النزعة المركزية وأكثرها دقة وتداولاً. وتوجد عدة أسس لتحديد قيمة المتوسط لمجموعة من البيانات مما أدى إلى وجود عدد من مقاييس المتوسط أهمها: المتوسط الحسابى، المتوسط الهندسى والمتوسط التوافقى.

(١) المتوسط الحسابى Arithmatic Mean

يعتبر المتوسط الحسابى من أشهر مقاييس المتوسطات وأسهلها حساباً وأكثرها

استخداماً فى عملية وصف البيانات. فهو يلخصها فى قيمة محددة تتركز حولها أغلب المفردات، فتمثل لها اتجاه عام وتعتبر بالتالى من أحد معالم ومميزات هذه البيانات. والمتوسط الحسابى مهم جداً فى الدراسات الجغرافية لأنه يعطى فكرة عامة عن الظاهرة موضع الدراسة. كأن نعرف متوسط سقوط الأمطار فى منطقة ما، أو متوسط إنتاج الفدان لمحصول من المحاصيل الزراعية، أو متوسط دخل الفرد من الدخل القومى فى بلد ومقارنة ذلك ببلدان أخرى.

ومن المعروف أن المتوسط الحسابى لمجتمع ظاهرة ما يعد قيمة ثابتة، بينما تعد قيمة المتوسط الحسابى لعينة قيمة متغيرة تختلف من عينة إلى أخرى للمجتمع الواحد. فمثلاً إذا سحبنا عدداً من العينات من مجتمع ما فإن المتوسط الحسابى غالباً ما يختلف من عينة إلى أخرى. ولو أن المتوسط الحسابى لمتوسطات هذه العينات يمكن اعتباره تقدير لمتوسط المجتمع. وفى معظم الحالات يعتبر المتوسط الحسابى للعينة قيمة غير متحيزة Unbiased لمتوسط المجتمع. أو بمعنى آخر أن المتوسط الحسابى لأى عينة قد يزيد أو يقل عن متوسط المجتمع الذى سحبت منه العينة إلا أن متوسط متوسطات العينات يطابق المتوسط العام للمجتمع. وسنعود لدراسة هذا الموضوع بالتفصيل فيما بعد.

المتوسط الحسابى للبيانات غير المبوبة:

يمكن تعريف المتوسط الحسابى لمجموعة من المفردات على أساس أنه القيمة التى لو أعطيت لكل مفردة من مفردات البيانات لكان مجموع هذه القيم الجديدة هو نفس مجموع المفردات الأصلية. كما يعرف المتوسط الحسابى (حسابياً) على أساس أنه القيمة الناتجة من جمع قيمك المفردات كلها مقسوماً على عدد المفردات. فمثلاً لو كان لدينا مجموعة (ن) من المفردات $س_1$ ، $س_2$ ، $س_3$ ... إلخ، فإن متوسطها الحسابى والذى يرمز له ($\bar{س}$) يمكن حسابه كالآتى:

$$\bar{س} = \frac{\text{مجموع قيم المفردات}}{\text{عدد المفردات}} = \frac{س_1 + س_2 + س_3 + س_ن}{ن}$$

$$\bar{س} = \frac{مجـ\ س}{ن} \quad \ldots \quad \ldots\ldots \quad \ldots\ldots \quad (٤ - ١)$$

والمعادلة الجبرية السابقة يمكن تطبيقها في كل الحالات التى تكون فيها البيانات ذات قيم مفردة. لذا فإنها تسمى بالأسلوب المباشر لحساب المتوسط الحسابي.

مثال تطبيقي:

لإيجاد المتوسط السنوى للأمطار فى كل من مدينتي سيدنى Sydney (استراليا) وستانلى مور Stanley Moor (انجلترا) خلال فترة ٢٠ سنة الموضحة بالجدول رقم (٤-١) تجرى الخطوات التالية.

١- تجمع كمية المطر (س) العشرين لكل مدينة على حدة.

٢- يقسم المجموع الكلى على عدد السنين (ن) لكل من المدينتين.

حدول رقم (٤-١)

كمية المطر السنوى (سنتيمتر) فى سيدنى (استراليا)

وستانلى مور (انجلترا) فى الفترة من ١٩٣٣ - ١٩٥٢

السنة	سيدنى	ستانلى مور	السنة	سيدنى	ستانلى مور
١٩٣٣	١٠٨	١٠٦	٤٤	٧٩	١٠٤
٣٤	١٦٥	١٣٨	٤٥	١٨٠	١٤٤
٣٥	٧٩	١٢٥	٤٦	٩٢	١٠٨
٣٦	٧٧	١٠٣	٤٧	١٠٥	١٥٢
٣٧	١٣٢	١٢٨	٤٨	٩٩	١١٩
٣٨	٩٩	١٣٣	٤٩	١٦٨	١٣٥
٣٩	٨٥	١١٨	٥٠	٢١٩	١٥٥
٤٠	١٠٠	١١٧	٥١	١٣٥	١٣٤
٤١	٦٨	١٢٠	١٩٥٢	١٥٠	١١٦
٤٢	١٢٣	١١٤			
٤٣	١٢٩	١٣٠	المجموع	٢٣٩٢	٢٤٩٨

وعلى ذلك يكون المتوسط الحسابى للمطر فى كل مدينة هو:

$$\bar{س}\ (\text{سيدنى}) = \frac{\text{مجـ س}}{\text{ن}} = \frac{٢٣٩٢}{٢٠} = ١١٩,٦\ \text{سنتيمتر}$$

$$\bar{س}\ (\text{ستانلى مور}) = \frac{\text{مجـ س}}{\text{ن}} = \frac{٢٤٩٨}{٢٠} = ١٢٤,٩\ \text{سنتيمتر}$$

وهذا يفسر جغرافيا على أنه على الرغم من اختلاف كمية المطر السنوية من مدينة لأخرى إلا أنه يوجد اتجاه عام وهو أن متوسط كمية المطر خلال فترة العشرين عاماً هى ١١٩,٦ سنتيمتر فى سيدنى و ١٢٤,٩ سنتيمتر فى ستانلى مور. ويعد هذا المؤشر من خصائص مجتمع المطر فى كل مدينة من المدينتين.

وفى حالة وجود عدد كبير من المفردات لاينصح باستخدام الأسلوب المباشر فى قياس المتوسط الحسابى حيث يطول وقت العمليات الحسابية والتى بالتالى تكون عرضه لاحتمالات الخطأ. ولكن ينصح باستخدام أسلوباً مختصراً لقياس المتوسط الحسابى يعتمد على اختصار الوقت وتبسيط العمليات الحسابية وتقليل درجة الأخطاء الممكنة، مع الحصول على نفس النتائج التى يمكن التوصل إليها باتباع الأسلوب المباشر. ويستعين الأسلوب المختصر بإحدى الطرق التالية لتحديد المتوسط الحسابى: طريقة العامل المشترك أو طريقة الوسط الفرضى. وفيما يلى خطوات إجراء كل منهما.

١ - طريقة العامل المشترك:

تعتمد هذه الطريقة على قسمة القيم الأصلية للمفردات على مقدار ثابت يطلق عليه اسم «العامل المشترك» وذلك لتبسيط العمليات الحسابية. وتقوم فكرة العامل المشترك على أساس أن المتوسط الحسابى لمجموعة من مفردات البيانات يمكن الحصول عليه لو أننا ضربنا ناتج قسمة مجموع المفردات المختصرة (أى بعد قسمتها على العامل المشترك) على عددها فى هذا العامل المشترك. فمثلاً إذا كانت لدينا المفردات التالية:

$س_١$، $س_٢$ ، $س_٣$، ....، $س_ن$

وعلى فرض أن (ل) تمثل العامل المشترك بين هذه المفردات، فإن قيم المفردات المختصرة (المفردات الأصلية مقسومة على العامل المشترك) تصبح كالآتى:

$$\frac{س_١}{ل} ، \frac{س_٢}{ل} ، \frac{س_٣}{ل} .... ، \frac{س_ن}{ل}$$

وعل ذلك فإننا نحصل على المتوسط الحسابى بواسطة الصيغة التالية:

$$\text{المتوسط الحسابى} = \frac{\text{مجموع قيم المفردات المختصرة}}{\text{عددها}} \times \text{العامل المشترك}$$

$$\bar{س} = \frac{مجـ \frac{س}{ل}}{ن} \times ل \text{ .... } (٤-٢)$$

وهناك شرط يجب مراعاته عند أخذ العامل المشترك لمجموعة من المفردات وهو أن يمثل العامل المشترك مقدار تسهل معه العمليات الحسابية ولا يسبب وجود كسور عشرية فى ناتج قسمة المفردات الأصلية.

مثال تطبيقى:

إذا كانت لدينا بيانات مجموعة من عشرة أشخاص عمر كل منهم على النحو التالى:

٢٠، ٣٠، ١٠، ٤٠، ٣٠، ٢٠، ٥٠، ٤٠، ٣٠، ١٠

والمطلوب: إيجاد متوسط السن لهذه المجموعة بإستخدام طريقة العامل المشترك فإننا نتتبع الخطوات الآتية:

١- نختار عاملاً مشتركاً هو القيمة (١٠) فى توزيع أعمار الأشخاص العشرة.

٢- نقسم مفردات البيانات (قيم الأعمار) على هذا العامل المشترك فنحصل بذلك على المفردات المختصرة للبيانات.

٣- تجمع القيم المختصرة ونقسم على عدد المفردات ثم يضرب الناتج فى العامل المشترك حتى نحصل على المتوسط الحسابى.

المتوسط الحسابى $\bar{\text{س}} = \frac{٢٠}{١٠} + \frac{٣٠}{١٠} + \frac{١٠}{١٠} + \frac{٤٠}{١٠} + \frac{٣٠}{١٠} + \frac{٢٠}{١٠}$

$+ \frac{٥٠}{١٠} + \frac{٤٠}{١٠} + \frac{٣٠}{١٠} + \frac{١٠}{١٠}$

$= \left[(٢ + ٣ + ١ + ٤ + ٣ + ٢ + ٥ + ٤ + ٣ + ١) \div ١٠ \right] \times ١٠$

$\therefore \bar{\text{س}} = \left(\frac{٢٨}{١٠} \right) \times ١٠ = ٢,٨ \times ١٠ = ٢٨$ سنة

٢- طريقة الوسط الفرضى:

تستخدم هذه الطريقة لتسهيل وإختصار العمليات الحسابية لإيجاد المتوسط الحسابى لعدد المفردات. وتقوم فكرة هذه الطريقة على أساس اختيار أى قيمة من بين قيم المفردات كوسط فرضى ثم نوجد إنحرافات قيم المفردات جميعها عن الوسط الفرضى المختار. وبعد ذلك نوجد متوسط الإنحرافات عن الوسط الفرضى ويضاف ذلك إلى الوسط الفرضى نفسه الذى يرمز له بالرمز (أ)، فنحصل على المتوسط الحسابى الحقيقى للمفردات، ويتم ذلك بالصيغة الإحصائية الآتية:

$$\text{المتوسط الحسابى} = \text{الوسط الفرضى} + \frac{\text{مجموع إنحرافات قيم المفردات عن وسطها الفرضى}}{\text{عدد المفردات}}$$

فعلى فرض أنه لدينا مجموع المفردات ($\text{س}_{١} + \text{س}_{٢}$ ، $\text{س}_{٣}$، $\text{س}_{\text{ن}}$) وعلى فرض أن (أ) هى الوسط الفرضى لهذه المجموعة، فإن الإنحرافات عنه تأخذ الشكل التالى:

$(\text{س}_{١} - \text{أ})$ ، $(\text{س}_{٢} - \text{أ})$ ، $(\text{س}_{٣} - \text{أ})$ ، $(\text{س}_{\text{ن}} - \text{أ})$

أى أن المتوسط الحسابى =

$$\text{أ} + \frac{(\text{س}_{١} - \text{أ}) + (\text{س}_{٢} - \text{أ}) + (\text{س}_{٣} - \text{أ}) + \ldots\ldots + (\text{س}_{\text{ن}} - \text{أ})}{\text{ن}}$$

وعلى ذلك فإن :

$$\bar{س} = أ + \frac{مجـ (س - أ)}{ن}$$

وبإعطاء الرمز (ح) للتعبير عن إنحرافات قيم المفردات عن وسطها الفرضى (س - أ) فإن

$$\bar{س} = أ + \frac{مجـ ح}{ن} \quad ...... \quad .... \quad (٤ -٣)$$

وعند إستخدام طريقة الوسط الفرضى كأسلوب احصائى مختصر لإيجاد المتوسط الحسابى للبيانات فإنه يجب علينا مراعاة مايلى:

أ- أن لا يكون الوسط الفرضى أصغر أو أكبر من قيم المفردات حتي لا يصبح مجموع الانحرافات دائماً موجباً أو سالباً بل يفضل أن تكون قيمة الوسط الفرضى هى القيمة التى تجعل مجموع الانحرافات أقل ما يمكن.

ب- يفضل أخذ الوسط الفرضى من أحد قيم المفردات الأكثر ترددا أو تكراراً أو تلك التى تتوسط عدد المفردات من حيث القيمة.

مثال تطبيقى:

إذا اعتبرنا أن الوسط الفرضى هو (٣٠) فى توزيع أعمار الأشخاص العشرة السابق ذكره، فإن إيجاد قيمة المتوسط الحسابى لهذا التوزيع بطريقة الوسط الفرضى يعتمد على تنظيم البيانات حسب الخطوات السابقة ووضعها على شكل جدول، كما هى الحال فى الجدول التالى:

جدول رقم (٤–٢)

حساب المتوسط الحسابى بطريقة الوسط الفرضى

رقم الشخص	السن (س)	(س – أ)	(ح)
١	٢٠	٢٠ – ٣٠	– ١٠
٢	٣٠	٣٠ – ٣٠	صفر
٣	١٠	١٠ – ٣٠	– ٢٠
٤	٤٠	٤٠ – ٣٠	١٠
٥	٣٠	٣٠ – ٢٠	صفر
٦	٢٠	٢٠ – ٣٠	– ١٠
٧	٥٠	٥٠ – ٣٠	٢٠
٨	٤٠	٤٠ – ٣٠	١٠
٩	٣٠	٣٠ – ٣٠	صفر
١٠	١٠	١٠ – ٣٠	– ٢٠
			+ ٤٠
			– ٦٠
			–٢٠

وعلى ذلك فإن

متوسط عمر الأشخاص = ٣٠ + $\left(\frac{-٢٠}{١٠}\right)$ = ٢٨ سنة

ومن الملاحظ أننا حصلنا على نفس الجواب السابق ولكن باستخدام أسلوب حسابى مبسط.

المتوسط الحسابى للبيانات المبوبة

لا تختلف عملية حساب المتوسط الحسابى كثيراً إذا كانت قيم المفردات تحدث بتكرارات، أى من نوع البيانات المبوبة، عنها مع البيانات غير المبوبة. غير أنع فى حالة التوزيعات التكرارية البسيطة للبيانات يجب ضرب كل قيمة فى تكرارها المناظر حتى نحصل على مجموع القيم ويقسم الناتج على المجموع الكلى للتكرارات. ولنفرض أن لدينا توزيعاً تكرارياً يشتمل على:

قيم المفردات: س_1 ، س_2 ، س_3 .... ، $\text{س}_\text{ن}$		المجموع
التكرار : ك_1 ، ك_2 ، ك_3 .... ، $\text{ك}_\text{ن}$		مج ك

فيكون المتوسط الحسابى هو = $\dfrac{\text{مجموع (القيمة × التكرار)}}{\text{مجموع التكرارات}}$

$$= \frac{\text{س}_1 \text{ك}_1 + \text{س}_2 \text{ك}_2 + \text{س}_3 \text{ك}_3 \ldots\ldots + \text{س}_\text{ن} \text{ك}_\text{ن}}{\text{ك}_1 + \text{ك}_2 + \text{ك}_3 + \ldots\ldots + \text{ك}_\text{ن}}$$

وعلى ذلك فإن:

$$\bar{\text{س}} = \frac{\text{مجـ س ك}}{\text{مجـ ك}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(٤-٤)}$$

ويطلق فى بعض الأحيان على التكرارات اسم معاملات الترجيح أو أوزان، وهذه تعتمد على الدلالة أو الأهمية المرتبطة بقيم المفردات حتى أن المتوسط الحسابى الذى نحصل عليه بالمعادلة (٤-٤) يعرف باسم المتوسط الحسابى المرجح (الموزون) Weighted Arithmatic Mean

مثال:

الجدول الآتى يبين توزيع حجم ٢٠ أسرة والمطلوب إيجاد متوسط حجم هذه الأسر.

جدول رقم (٤-٣)

حساب المتوسط الحسابى المرجح

حجم الأسرة	عدد الأسر (التكرار) ك	حجم الأسرة × التكرار س × ك
١	٣	٣
٢	٤	٨
٣	٢	٦
٤	٤	١٦
٥	٢	١٠
٦	٢	١٢
٧	٢	١٤
٩	١	٩
المجموع ٣٧	٢٠	٧٨

وبتطبيق الصيغة الرياضية (٤-٤) فإن:

متوسط حجم الأسرة المرجح (سَ) = $\frac{٧٨}{٢٠}$ = ٣,٩ (أى ٤ أشخاص تقريباً

بينما يكون المتوسط الحسابى العادى هو:

سَ = $\frac{٣٧}{٨}$ = ٤,٦٢٥ (أى ٥ أشخاص تقريبا)

وهنا يظهر أن المتوسط الحسابى العادى (٤,٦٢٥) يعد متوسط لايمثل مركز البيانات الحقيقى. بينما نستنتج أن المتوسط الحسابى المرجح (٣,٩) أكثر ملائمة فى هذه الحالة لاختلاف الأهمية أو الدقة النسبية للبيانات ممثلة فى التكرارات المناظرة لقيم المفردات (حجم الأسر).

ولتبسيط عملية الحساب يمكن أن نستخدم وسط فرضى ونوجد انحرافات قيم المفردات عن هذا الوسط (كما سبق شرحه)، ثم نضرب الإنحرافات فى التكرارات (الأوزان) المناظرة وأخيراً. يؤخذ متوسط مجموع الإنحرَافات ويضاف على الوسط الفرضى فنحصل بذلك على المتوسط الحسابى المرجح.

المتوسط الحسابى (س)

$$= ١ + \frac{(\text{س}_{١} - \text{أ})\,\text{ك}_{١} + (\text{س}_{٢} - \text{أ})\,\text{ك}_{٢} - (\text{س}_{٣} - \text{أ})\,\text{ك}_{٣} + \dots + (\text{س}_{\text{ن}} - \text{أ})\,\text{ك}_{\text{ن}}}{\text{ك}_{١} + \text{ك}_{٢} + \text{ك}_{٣} \dots\dots + \text{ك}_{\text{ن}}}$$

$$\bar{\text{س}} = ١ \div \left(\frac{\text{مجـ}\,(\text{س} - \text{أ})\,\text{ك}}{\text{مجـ}\,\text{ك}} \right)$$

وبالتعويض عن (س − ١) بالرمز ح فإن

$$\bar{\text{س}} = \text{أ} + \left(\frac{\text{مجـ}\,\text{ح}\,\text{ك}}{\text{مجـ}\,\text{ك}} \right) \dots\dots\dots\dots\dots\dots (٤-٥)$$

فإذا اعتبرنا أن الوسط الفرضى (أ) للمثال السابق، هو القيمة (٤) فإنه يمكن الحصول على الجدول الآتى:

جدول رقم (٤-٤)

حساب المتوسط الحسابى المرجح بطريقة الوسط الفرضى

حجم الأسرة س	عدد الأسر ك	ح = س - أ	ح ك
١	٣	-٣	-٩
٢	٤	-٢	-٨
٣	٢	-١	-٢
٤	٤	صفر	صفر
٥	٢	١	٢
٦	٢	٢	٤
٧	٢	٣	٦
٩	١	٥	٥
المجموع	٢٠		-١٩ +١٧

أى المتوسط الحسابى = الوسط الفرضى + $\frac{\text{مجـ ح ك}}{\text{مجـ ك}}$

= ٤ + $\left[\frac{-٢}{٢٠}\right]$

= ٤ - ٠,١

= ٣,٩ وهى نفس النتيجة السابقة

المتوسط الحسابى لجداول التوزيعات التكرارية:

سبق القول أن لاختصار عدد كبير من المفردات فإننا نحولها إلى فئات ونسجل عدد قيم المفردات (التكرارات) فى كل فئة فنحصل على جدول تكرارى.

ولاشك أن حساب المتوسط الحسابى من مثل هذا الجدول يكون أسهل وأسرع عن جمع عدد القيم كلها. ولما كان المتوسط الحسابى يعتمد فى قياسه على مجموع قيم المفردات فإن وجود القيم فى شكل فئات ينفى حقيقة القيم ويظهرها فى شكل مقادير تنحصر بين حدى الفئة مما يجعل من الصعب تحديد مجموع القيم تحديداً دقيقاً. لذلك فإنه للحصول على حقية القيم نفترض أن كل عدد معين من التكرارات يحدث فى منتصف مدى الفئة وأن القيم تتركز فى مراكز فئاتها. وعلى ذلك فإن الخطوة الأولى فى حساب المتوسط الحسابى لتوزيع تكرارى (حسب الفئات) هى إيجاد مراكز فئات التوزيع كمايلى:

$$\text{مركز الفئة (م)} = \frac{\text{الحد الأعلى للفئة + الحد الأدنى للفئة}}{٢}$$

ويتحويل الفئات إلى مراكز الفئات (م) فإننا نعتبر الأخيرة هى القيم التى نريد إيجاد متوسطها الحسابى، علما بأن لكل منها تكراراً معيناً (مجموع القيم يساوى مجموع التكرارات كلها) وبذلك تكون خطوات العمل لحساب المتوسط الحسابى كمايلى:

١- تجميع القيم فى فئات مناسبة وتحديد عدد مرات حدوث كل قيمة (التكرارات).

٢- تحديد مراكز الفئات (م).

٣- اختيار وسط فرضى (أ) من بين مراكز الفئات. ويستحسن أن يكون مركز الفئة التى تقابل أكبر تكرار.

٤- حساب انحرافات (ح) مراكز الفئات عن الوسط الفرضى. أى ح = (م - أ).

٥-ضرب كل انحرافات فى التكرار المناظر له (ح × ك) ثم إيجاد حاصل الضرب مجـ (ح ك).

٦- قسمة المجموع الناتج فى الخطوة ٤ على مجموع التكرارات ($\frac{\text{مجـ ح ك}}{\text{مجـ ك}}$).

٧- إضافة ناتج خارج القسمة إلى الوسط الفرضى (أ) فنحصل على المتوسط الحسابى المطلوب.

٨- يمكن أن نقسم الانحرافات (خطوة ٤) على طول الفئة (ل) إذا كان التوزيع منتظماً (متساوياً) فى فئاته. فنحصل على $\frac{ح}{ل}$ ويكون المتوسط الحسابى فى هذه الحالة هو:

$$\bar{س} = أ + \left[\frac{مجـ ح \times ك}{مجـ ك}\right] \times ل \qquad .... (٤-٦)$$

وتسمى الصيغة الرياضية السابقة بطريقة الترميز Coding Method عند حساب المتوسط الحسابى. وهذه الطريقة مختصرة جداً ويجب استخدامها دائماً للبيانات التكرارية (المجمعة) عندما تكون أطوال الفئات متساوية. فإذا كانت لدينا بيانات عن توزيع مساحات المزارع فى منطقة ما (بالفدان) والبالغ عددها ٧٨ مزرعة مثلاً وأردنا حساب المتوسط المساحى لها فستكون خطوات الحساب على النحو الموضح بالجدول (٤-٥).

جدول رقم (٤-٥)

توزيع مساحات المزارع فى منطقة ما بالفدان

الفئات مساحة المزارع	التكرار ك عدد المزارع	مراكز الفئات م	الإنحراف عن الوسط الفرضى ل/ك	الإنحراف × التكرار ل/ح × ك
١ - ١٠	١	٥,٥	-٦	-٦
١١ - ٢٠	٣	١٥,٥	-٥	-١٥
٢١ - ٣٠	٤	٢٥,٥	-٤	-١٦
٣١ - ٤٠	٦	٣٥,٥	-٣	-١٨
٤١ - ٥٠	٩	٤٥,٥	-٢	-١٨
٥١ - ٦٠	٩	٥٥,٥	-١	-٩
٦١ - ٧٠	١٠	٦٥,٥	صفر	صفر
٧١ - ٨٠	٩	٧٥,٥	١	٩
٨١ - ٩٠	٨	٨٥,٥	٢	١٦
٩١ - ١٠٠	٦	٩٥,٥	٣	١٨
١٠١ - ١١٠	٤	١٠٥,٥	٤	١٦
١١١ - ١٢٠	٣	١١٥,٥	٥	١٥
١٢١ - ١٣٠	٣	١٢٥,٥	٦	١٨
١٣١ - ١٤٠	٢	١٣٥,٥	٧	١٤
١٤١ - ١٥٠	١	١٤٥,٥	٨	٨
المجموع	٧٨			- ٨٢ + ١١٤

وباتباع الخطوات السابقة لإيجاد المتوسط الحسابى نلاحظ الآتى:

الوسط الفرضى للتوزيع = ٦٥,٥

مجمـــــوع التكرارات = ٧٨

مجمـــــوع الإنحرافات = + ٣٢

وحيث أنه يمكن إيجاد المتوسط بتطبيق المعادلة (٤-٦)، وبالتعويض من حسابات الجدول فإن:

$$س َ = ٦٥,٥ + \left(\frac{٣٢}{٧٨} \times ١٠ \right)$$

$$= ٦٥,٥ + ٤,١$$

$$= ٦٩,٨ \text{ فداناً}$$

٢- المتوسط الهندسي Geometeric Mean

تتميز العلوم الإجتماعية، ومن بينها الجغرافيا، بأن بعض ظاهراتها لا تتغير في نسق منتظم أو بمقادير متساوية من فترة لأخرى، بل أنها تتغير بمعدلات مختلفة ومتباينة. ففي الديموجرافيا (علم السكان) تؤكد الدراسة على أن التغير السكاني لا يأخذ شكل المتوالية العددية، أي أن السكان لا يتزايدون أو يتناقصون بعدد متساوي من فترة لأخرى، بل أن التغير السكاني لا يتزايدون أو يتناقضون بعدد متساوي من فترة لأخرى، بل أن التغير السكاني يأخذ شكل المتوالية الهندسية، أي بنسبة تتناسب مع عدد السكان كل سنة وبناء على ذلك فإنه لا يمكن استخدام المتوسط الحسابي في قياس التغير السكاني باعتباره مجموع قيم المفردات على عددها، بل يجب استخدام الجذر التوني لحاصل ضرب أعداد السكان لفترات من السنين وهو ما يعرف باسم المتوسط الهندسي. ومن ثم فإن المتوسط الهندسي الذي يرمز له بالرمز (هـَ)، لمجموعة من المفردات (ذات قيم موجبة) عددها (ن) يساوي الجذر النوني لحاصل ضرب قيم هذه المفردات. فعلى فرض أن لدينا عدد من المفردات على النحو التالي:

$س_١$، $س_٢$، $س_٣$، ....، $س_ن$ فيكون متوسطها الهندسي هو :

$$\text{المتوسط الهندسي (هـ)} = \sqrt[ن]{س_١ \times س_٢ \times س_٣ \times ...... \times س_ن}$$

ولتسهيل إيجاد قيمة (هـ) نستعين باللوغارتيمات فيكون:

$$\text{لو هـ} = \frac{1}{\text{ن}} \left(\text{لو س}_{1} + \text{لو س}_{2} + \text{لو س}_{3} + \ldots\ldots \text{لو س}_{\text{ن}}\right)$$

$$\text{لو هـ} = \frac{\text{مجـ لو س}}{\text{ن}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (\text{٤} - \text{٧})$$

بمعنى أن لوغاريتم المتوسط الهندسى عبارة عن المتوسط الحسابى للوغارتيمات القيم المستخدمة فى حساب المتوسط الهندسى نتبع ماسبق شرحه - سابقاً - فى إيجاد المتوسط الحسابى مستخدمين لوغاريتمات قيم المفردات. ومن الناحية العملية فإن المتوسط الهندسى أقل تأثراً فى حسابه بالقيم المتطرفة عن المتوسط الحسابى. ومن ثم فإن المتوسط الهندسى لمجموعة من مفردات البيانات يكون أقل من المتوسط الحسابى لنفس المفردات مالم تكن قيم المفردات جميعها متساوية. فمثلاً إذا كانت لدينا قيم المفردات الآتية والتى تمثل أطوال بعض الطرق بالكيلو مترات وفى منطقة ما:

١٢٣، ١٢٠، ١٣٨، ١٦٠، ١٤٠، ١٧٢، ٢٥٠

وأردنا حساب المتوسط الهندسى فإننا نوجد لوغاريتمات هذه القيم وهى على الترتيب:

٢,٠٨٩٩، ٢,١١٣٩، ٢,١٣٩٩ ٢,٢٠٤١، ٢,١٤٦١، ٢,٢٣٥٥، ٢,٣٩٧٩

وبتطبيق المعادلة ٤٠-٧) فإن :

$$\text{لو هـ} = \frac{\text{٢,٠٨٩٩} + \text{٢,١١٣٩} + \text{٢,١٣٩٩} + \text{٢,٢٠٤١} + \text{٢,١٤٦١} + \text{٢,٢٣٥٥} + \text{٢,٣٩٧٩}}{\text{٧}}$$

$$= \text{٢,١٨٦٧}$$

وبالكشف فى جداول الأعداد المقابلة اللوغاريتمات (الرقم ٢,٥٨٦٧) نجد أن المتوسط الهندسى:

$$\text{هـ} = \text{١٥٣,٨١}$$

أما إذا قمنا بحساب المتوسط الحسابى لنفس القيم بالمعادلة (٤-٦) نجد أن:

$\bar{س} = \frac{١١١٣}{٧} = ١٥٩$ (وهو كما ذكرنا أكبر من المتوسط الهندسى).

ولحساب المتوسط الهندسى من جدول التوزيع التكرارى للبيانات فإننا نتبع نفس الأسلوب الذى استخدمناه لحساب المتوسط الحسابى بعد أن نوجد لوغاريتمات مراكز الفئات. ثم حساب المتوسط الهندسى كمايلى:

$$\text{لو } \bar{هـ} = \frac{\text{مجـ ك لو س}}{\text{مجـ ك}} \qquad .... \qquad (٤-٨)$$

مثال: الجدول التالى يمثل توزيع نسب أسعار ١٠٠ سلعة والمطلوب حساب متوسطها الهندسى.

جدول رقم (٤-٦)

حساب المتوسط الهندسى لنسب أسعار ١٠٠ سلعة

الفئات	التكرار	مراكز الفئات	لو م	ك × لو م
١١٥-	١٠	١٢٠	٢,٠٧٩٢	٢٠,٧٩٢
١٢٥-	٢٠	١٣٠	٢,١١٣٩	٤٢,٢٧٨
١٣٥-	٤٠	١٤٠	٢,١٤٦١	٨٥,٨٤٤
١٤٥-	٢٠	١٥٠	٢,١٧٦١	٤٣,٥٢٢
١٥٥-	١٠	١٦٠	٣,٢٠٤١	٢٣,٠٤٢
المجموع	١٠٠			٢١٤,٤٧٧

وعلى ذلك فإن:

$$\text{لو } \bar{هـ} = \frac{\text{مجـ (ك × لو س)}}{\text{مجـ ك}} = \frac{٢١٤,٤٧٧}{١٠٠} = ٢,١٤٤٧٧$$

ويكون المتوسط الهندسى لهذا التوزيع = ١٣٩,٦

ويعد المتوسط الهندسى من أنسب وأصلح مقاييس المتوسطات التى تصف الإتجاه العام لمجموعة من النسب أو المعدلات وبصفة خاصة معدلات النمو، ومعدلات المواليد كما فى المثال التالى.

مثال : فيمايلى معدل النمو السنوى للسكان لكل عشر سنوات فى مدينة ما والمطلوب حساب متوسط معدل النمو فى عدد السكان.

السنة	١٩٣٠	١٩٤٠	١٩٥٠	١٩٦٠	١٩٧٠	١٩٨٠
معدل النمو	٣,٥	٢,٥	٢,٠	١,٥	١,٠	٠,٥

∴ متوسط معدل النمو = $\sqrt[٦]{٣,٥ \times ٢,٥ \times ٢,٠ \times ١,٥ \times ١,٠ \times ٠,٥}$

لو هـَ = $\frac{١}{٦}$ (لو ٣,٥ + لو ٢,٥ + لو ٢,٠ + لو ١,٥ + لو ١,٠ +٠,٥)

= $\frac{١}{٦}$ ×

(٠,٥٤٤١ + ٠,٣٨٧٩ + ٠,٣٠١٠ + ٠,١٧٦١ + ٠,٠٠ + ١,٦٩٨٩)

= $\frac{١}{٦}$ × ١,١١٨٠ = ٠,١٨٦٣

هـَ = ١,٥٣٥٧

كذلك يصلح المتوسط الهندسى لقياس تقدير عدد السكان فى سنوات ما بين التعدادات. وهنا يكون التغير فى عدد السكان متناسباً مع عدد السكان نفسه.

مثال: كان سكان مصر فى تعداد سبتمبر ١٩٦٠ هو ٢٥٩٨٤٢٠١ نسمة وكان تعدادها فى مايو سنة ١٩٦٦ هو ٢٩٩٤٣٨١٠ نسمة. فما هو تعداد سكان مصر فى منتصف الفترة بين التعدادين أى فى ١٩٦٣ . فإذا حسبنا التعداد فى منتصف الفترة عن طريق المتوسط الحسابى فإنه يكون :

سَ = $\frac{٢٥٩٨٤٢٠١ + ٢٩٩٤٣٨١٠}{٢}$ = ٢٧٩٦٤٠٠٥ نسمة

ولكن يعنى هذا أن الزيادة فى عدد السكان كل عام يكون بقدر متساوى وهذا لا يكون صحيحاً لأن معدل النمو فى هذه الحالة يتزايد مع تزايد السكان، ولذا فإن التقدير الصحيح يتم بإستخراج المتوسط الحسابى الهندسى. وحيث أن عدد القيم هو ٢ فإن :

$$\acute{هـ} = \sqrt{(٢٥٩٨٤٢٠١ \times ٢٩٩٤٣٨١٠)}$$

$$\text{لو } \acute{هـ} = \frac{١}{٢}\ (\text{لو } ٢٥٩٨٤٢٠١ + \text{لو } ٢٩٩٤٣٨١٠)$$

$$\text{لو } \acute{هـ} = \frac{١}{٢}\ (٧,٤١٤٦٦ + ٧,٤٧٦٢٦)$$

$$\text{لو } \acute{هـ} = ٧,٤٤٥٤٦$$

$$\acute{هـ} = ٢٧٨٩٣٠٠٠ \text{ نسمة}$$

ويتضح من كل مما سبق أن المتوسط الهندسى على الرغم من مميزاته فى إستخراج المتوسطات للنسب والمعدلات إلا أنه أكثر صعوبة فى طريقة الحساب والفهم من المتوسط الحسابى.

٣- المتوسط التوافقى Harmonic Mean

يعرف المتوسط التوافقى لمجموعة من قيم المفردات بأنه مقلوب المتوسط لمقلوبات هذه القيم. فإذا فرضنا أنه كان لدينا القيم $س_١$، $س_٢$، $س_٣$، ...، $س_ن$ وعددها ن فإن متوسطها التوافقى هو :

$$ق = \frac{ن}{\frac{١}{س_١} + \frac{١}{س_٢} + \frac{١}{س_٣} + \ldots\ldots + \frac{١}{س_ن}} = \frac{ن}{مجـ \frac{١}{س}} \ldots (٤-٩)$$

فالمتوسط التوافقى للأعداد ٢٠٠، ٤٠٠، ٦٠٠، ٨٠٠، ١٠٠٠ هو :

$$\text{ق} = \frac{٥}{\frac{١}{٢٠٠} + \frac{١}{٤٠٠} + \frac{١}{٦٠٠} + \frac{١}{٨٠٠} + \frac{١}{١٠٠٠}} = \frac{٥}{٠,١١٤٢}$$

$$= ٤٣٧,٨٣$$

أما إذا حسبنا المتوسط الحسابى لهذه القيم فإنه يكون

$$\text{س} = \frac{٢٠٠ + ٤٠٠ + ٦٠٠ + ٨٠٠ + ١٠٠٠}{٥} = \frac{٢٠٠٠}{٥} = ٦٠٠$$

ويعتبر المتوسط التوافقى أنسب مقاييس المتوسط لتمثيل الأثمان ومعدلات السرعة وتعطى مثالاً يوضح تفسير إستخدام المتوسط التوافقى بدلاً من المتوسط الحسابى على النحو التالى :

مثال: لنفرض أن معدل سرعة الحركة على أحد الطرق الملاحية (٤٠٠ كيلو مترا) يختلف من جزء إلى آخر على الطريق، فيمكن لأى باخرة أن تقطع ١٠٠ ك.م الأولى بسرعة ٢٠ كيلومترا فى الساعة ، ثم ١٠٠ ك.م الثانية بسرعة ٣٠ كيلو مترا فى الساعة، ثم ١٠٠ ك.م الثالثة بسرعة ٤٠ كيلو متراً فى الساعة ثم ١٠٠ ك.م الأخيرة بسرعة ٥٠ كيلو متراً فى الساعة. فإذا حسبنا المتوسط الحسابى لسرعة الباخرة فإنه يكون :

$$\bar{\text{س}} = \frac{٢٠ + ٣٠ + ٤٠ + ٥٠}{٤} = \frac{١٤٠}{٤} = ٣٥ \text{ كيلومتراً / ساعة}$$

وهذا غير صحيح، أو أن المتوسط الحسابى يكون فى هذه الحالة مضللاً إلى حد كبير، وذلك لأن الباخرة قطعت المسافة الأولى فى خمس ساعات والمسافة الثانية فى ثلاث ساعات وثلث والمسافة الثالثة فى ساعتين ونصف والمسافة الرابعة فى ساعتين. أى أنها استغرقت وقتاً وقدره ١٢ ساعة و ٥٠ دقيقة. وعلى ذلك يكون متوسط السرعة هو :

متوسط السرعة = $\frac{٤٠٠}{١٢٨٣}$ = ٣١,١٧ كيلو مترا/ ساعة

والمتوسط السابق هو نفسه المتوسط التوافقى لسرعات الباخرة

$$ق = \frac{٤}{\frac{١}{٢٠} + \frac{١}{٣٠} + \frac{١}{٤٠} + \frac{١}{٥٠}} = \frac{٤}{\frac{٧٧}{١٠٠}}$$

= ٣١,١٧ كيلومترا/ ساعة

وللمتوسط التوافقى لمجموعة من القيم خاصية هامة وهى أنه أقل من كل من المتوسط الحسابى والمتوسط الهندسى مالم تكن جميع القيم متساوية. كما يمكن حساب المتوسط التوافقى لجداول التوزيعات التكرارية، وفيها يضاف عموداً جديداً إلى جدول التكرار يوضع فيه مقلوب مراكز الفئات ($\frac{١}{م}$) ، ثم نضرب هذه المقلوبات فى التكرارات المناظرة ($\frac{١}{م}$ × ك)، ثم يوجد حاصل الجمع [مجـ ($\frac{١}{م}$)] ، ثم يقسم عليه، مجموع التكرارات $\frac{مجـ ك}{مجـ \frac{ك}{م}}$.

مثال: الجدول الآتى يبين طريقة إيجاد المتوسط التوافقى لتوزيع تكرارى لسرعات التيارات المائية فى ١٠٠ رافد نهرى لأحد أحواض التصريف المائى

جدول رقم (٤-٧)

طريقة حساب المتوسط التوافقى لسرعة التيارات المائية ١٠٠ رافد نهرى

فئات السرعة سنتيمتر/ ثانية	عدد الأنهار ك	مراكز الفئات م	مقلوب مراكز الفئات $\frac{١}{م}$	ك × $\frac{١}{م}$
-١٢,٥	٢٠	١٥	٠,٠٦٦٠	١,٣٣٣
-١٧,٥	٥٠	٢٠	٠,٠٥٠٠	٢,٥٠٠
-٢٢,٥	٢٠	٢٥	٠,٠٤٠٠	٠,٨٠٠
-٢٧,٥	١٠	٣٠	٠,٠٣٣٣	٠,٣٣٣
المجموع	١٠٠			٤,٩٦٦

ويكون المتوسط التوافقى للسرعات هو:

$$ق = \frac{١٠٠}{٤,٩٦٦} = ٢٠,١٤ \text{ سنتيمتراً/ ثانية}$$

وعلى الرغم من الصعوبات التى تواجه حساب هذا المقياس وكذلك صعوبة فهم الاحصائية للمتوسط التوافقى إلا أنه يفضل على باقى المتوسطات فى مجال الجغرافية، إذا ما كانت المفردات موضع الدراسة منسوبة إلى ثابت معين مثل السرعة فى الساعة (أو الدقيقة أو الثانية) أو بصفة عامة معدلات التغير المنسوبة إلى أساس ثابت.

ثانياً: الوسيط Mediau

الوسيط هو المقياس الثانى من مقاييس النزعة المركزية. ويعرف الوسيط لمجموعة من المفردات أو مجموع من التكرارات بأنه القيمة التى تتوسط تلك المجموعة أو ذاك المجموع. أى أنه القيمة التى تقسم المجموعة إلى قسمين متساويين، القسم الأول منها يشمل عدد القيم الأصغر منها، والقسم الآخر يشمل عدد القيم الأكبر منها. ويمكن إيجاد الوسيط عن طريق ترتيب القيم تصاعدياً أو تنازلياً على هيئة منظومة Array للحصول على القيمة الوسطى لهذا الترتيب إذا كان عدد القيم فردياً أو متوسط القيمتين فى منتشف المنظومة إذا كان عدد القيم زوجياً.

فلو فرض أنه لدينا مجموعة المفردات التالية والتى تمثل أطوال سبعة أودية بالكيلو مترات ٨٣، ٦٨، ٦٦، ٦٥، ٧٣، ٦٩، ٧١. ثم نقوم بترتيب هذه القيم تصاعدياً (أو تنازلياً) فنحصل على:

٦٥ ، ٦٦ ، ٦٨ ، ٦٩ ، ٧١ ، ٧٣ ، ٨٣

ويكون الوادى الوسيط هو صاحب الطول الرابع فى الترتيب أى ٦٩ إذ أن هناك ثلاثة أطوال أكبر منه وثلاثة أطوال أصغر منه. ولإيجاد ترتيب الوسيط من مثل هذه المفردات ذات العدد الفردى فإنه يكون:

ترتيب الوسيط = $\frac{\text{عدد القيم} + 1}{2}$ = $\frac{ن + 1}{2}$ .... (٤-١٠)

أما إذا كان عدد قيم المفردات (ن) زوجياً فإن قيمة الوسيط (ط) تقع ما بين قيمة المفردة التى ترتبها ($\frac{ن}{2}$) وقيمة المفردة التى ترتيبها ($\frac{ن}{2} + 1$). أو بمعنى آخر أن ترتيب الوسيط يقع بين القيمتين المتوسطين فى الترتيب وأن قيمته يمكن الحصول عليها من حساب المتوسط الحسابى لهاتين القيمتين كما يبدو من المثال التالى لمعرفة القيمة الوسيطية لعمر ثمانية من الأشخاص (بالسنة):

٢٠، ٢١، ٢٦، ٢٧، ٢٩، ٣٠، ٣٢، ٣٣

وهنا نجد أن ترتيب الوسيط = $\frac{ن + 1}{2}$ = $\frac{8 + 1}{2}$ = ٤,٥

أى أن قيمة الوسيط تقع فى منتصف المسافة مابين المفردة الرابعة والخامسة أو قيمة الوسيط = $\frac{1}{2}$ (القيمة التى ترتيبها $\frac{8}{2}$ + القيمة التى ترتيبها ($\frac{8}{2}$+1))

= $\frac{1}{2}$ (القيمة الرابعة + القيمة الخامسة).

∴ قيمة الوسيط (ط) = $\frac{1}{2}$ (٢٧ + ٢٩) = ٢٨ سنة

إيجاد الوسيط للبيانات المبوبة:

لإيجاد الوسيط للبيانات المبوبة فى شكل جداول توزيعات تكرارية للفئات نتبع الخطوات التالية:

١ - ترتب التكرارات على شكل تكرارات متجمعة صاعدة أو هابطة.

٢ - يحدد ترتيب الوسيط بين التكرارات وهو = $\frac{\text{مجموع التكرارات}}{2}$

= $\frac{\text{مجـ ك}}{2}$ بصرف النظر ما إذا كانت «ك» فردية أو زوجية.

٣- يستخدم ترتيب الوسيط لتحديد الفئة التى يقع بها الوسيط (من عمود التكرار المتجمع بلاجدول) وتسمى الفئة الوسيطية. ثم تحدد الحدود الحقيقية للفئة الوسيطة وتكرارتها الأصلية.

٤- تحدد التكرارات لجميع الفئات التى تسبق الفئة الوسيطية بالاستعانة بالتوزيع التكرارى المتجمع الصاعد.

٥- يستخدم قانون استخراج الوسيط وهو:

الوسيط = الحد الأدنى الحقيقى للفئة الوسيطة +

$$\left[\left(\frac{\text{ترتيب الوسيط – التكرار المتجمع الصاعد السابق للفئة الوسيطية}}{\text{التكرار الأصلى للفئة الوسيطية}} \right) \times \text{طول الفئة} \right]$$

$$= \text{ف}_{\text{و}} + \left(\frac{\frac{1}{2}(\text{ن}) - \text{ك ص}}{\text{ك}_{\text{و}}} \right) \times \text{ل} \quad .... (٣ - ١١)$$

ولبيان طريقة حساب الوسيط للتوزيع التكرارى نستخدم بيانات الجدول التالى الذى يبين كميات الأمطار الساقطة بالمليمترات فى ٤٠ مرصداً جوياً فــى اقليم ما:

جدول رقم (٤ -٨)
حساب الكمية الوسيطية للأمطار الساقطة
فى ٤٠ مرصداً «بالملليمتر»

الأمطار ملليمترات	عدد المراصد التكرار	المتجمع الصاعد	
		أقل من الحد الأعلى للفئات	التكرار المتجمع الصاعد
١١٨ - ١٢٦	٣	أقل من ١٢٦	٣
١٢٧ - ١٣٥	٥	أقل من ١٣٥	٨
١٣٦ - ١٤٤	٩	إقل من ١٤٤	١٨
١٤٥ - ١٥٣	١٢	أقل من ١٥٣	٢٩ (فئة الوسيط)
١٥٤ - ١٦٢	٥	أقل من ١٦٢	٣٤
١٦٣ - ١٧١	٤	أقل من ١٧١	٣٨
١٧٢ - ١٨٠	٢	أقل من ١٨٠	٤٠
المجموع	٤٠		

وهناك طريقتان لإيجاد الوسيط مثل هذا التوزيع التكرارى هما:

(١) **الطريقة الأولى**، باستخدام الاستكمال نفترض أن كميات الأمطار فى الجدول التكرارى السابق (رقم ٤-٨) تتوزع توزيعاً مستمراً (متصلاً). فى هذه الحالة فإن الوسيط هو الكمية التى تقع نصف التكرارات الكلية أعلاها والنصف الآخر أسفلها وترتيبها هو نصف مجموع التكرارات ($\frac{٤٠}{٢}$ = ٢٠). وحيث أن مجموع تكرارات الفئات الثلاث الأولى (التكرار المتجمع الصاعد للفئات أقل من ١٤٤) هو ٣ + ٥ + ٩ = ١٧. وحى نحصل على الترتيب المطلوب ٢٠ فإننا نريد ٣ تكرارات من التكرارات الموجودة فى الفئة التالية وهى الفئة التى يقع فيها الوسيط (١٢ تكرار). وبما أن الفئة الرابعة أو الفئة الوسيطية (١٤٥ - ١٥٣) هى

فى الحقيقة تقابل الكميات ١٤٤.٥ - ١٥٣.٥ فإن الوسيط يقع فى ٣/١٢ من المسافة بين ١٤٤.٥ و ١٥٣.٥ أى أن الوسيط هو:

$$\text{الوسيط} = ١٤٤.٥ + \frac{٣}{١٢} (١٥٣.٥ - ١٤٤.٥)$$

$$= ١٤٤.٥ + \frac{٣}{١٢} (٩)$$

= ١٤٦.٨ ملليمتراً

(٢) **الطريقة الثانية**، باستخدام المعادلة (٤-١١). بما أن التكرارات المتجمعة الصاعدة للفئات الأولى والفئات الأربع الأولى هى على الترتيب ١٧، ٢٩ فإن الوسيط يقع فى الفئة الرابعة والتى هى بالتالى الفئة الوسيطة، وبذلك فإن:

الحد الأدنى الحقيقى للفئة الوسيطية ($ف_و$) = ١٤٤.٥

المجموع الكلى للتكرارات (ن) = ٤٠

مجموع التكرارات المتجمعة الصاعدة للفئات السبقة للفئة الوسيطية ($ك_ص$) = ١٧

التكرار الأصلى للفئة الوسيطة ($ك_و$) = ١٢

طـــــــول الفــــــــئة (ل) = ٩

وتكون قيمة الوسيط هى:

$$\text{الوسيط} = ف_و + \left(\frac{\frac{١}{٢}(ن) - ك_ص}{ك_و} \right) \times ل$$

$$= ١٤٤.٥ + \left(\frac{\frac{١}{٢}(٤٠) - ١٧}{١٢} \right) \times ٩$$

$$= ١٤٤.٥ + \left(\frac{٢٠ - ١٧}{١٢} \right) \times ٩$$

= ١٤٦.٨ ملليمتراً

وهى نفس النتيجة التى حصلنا عليها بطريقة الاستكمال.

تعين الوسيط بيانياً:

يختلف الوسيط عن المتوسط الحسابى فى أنه يمكن إيجاد قيمته عن طريق الرسم البيانى للمدرج التكرارى أو لمنحنيات التكرار المتجمع الصاعد أو الهابط. فلو فرض أننا نريد تعيين الوسيط بيانياً فى مثال كميات الأمطار السابق فإننا نقوم بتمثيل التوزيعات التكرارية فى الجدول رقم (٤-١) وذلك برسم مدرجها التكرارى أو المنحنى التكرارى المتجمع النسبى (الصاعد أو الهابط)، ثم نعين ترتيب الوسيط وقيمته على أى منها كمايلى:

فى الشكل رقم (٤-١) الذى يوضح المدرج التكرارى المقابل لكميات الأمطار فى المثال السابق، يمكن تعيين ترتيب الوسيط منه على أساس أن الوسيط هو الأحداثى السينى للخط هـ ط الذى يقسم المدرج التكرارى إلى مساحتين متساويتين. وحيث أن المساحة تقابل التكرار فى المدرج التكرارى، فإن الخط هـ ط يقسم المساحة الكلية بحيث تكون التكرارات على يمينه والتكرارات على يساره مساوية لنصف التكرارا الكلية أو ٢٠ كما أنه يقسم الفئة الوسيطية إلى قسمين تتناسب مساحتهما عكسياً مع مجموع التكرارات السابقة واللاحقة لترتيب الوسيط. فمثلاً المساحة ط هـ د ب تناظر التكرار ٣ الذى يضاف إلى التكرارات السابقة للوسيط وعددها ١٧ لتصبح ٢٠، والمساحة أ حـ هـ ط تناظر التكرار ٩ الذى يضاف إلى التكرارات اللاحقة للوسيط وعددها ١١ لتصبح ٢٠ أيضاً وبهذا فإن:

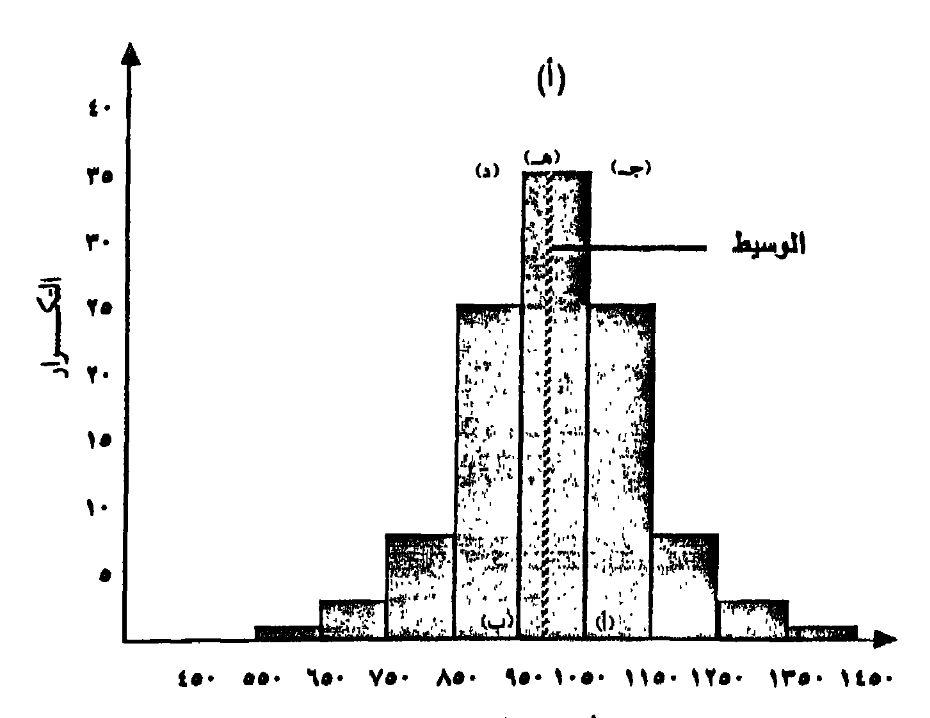
$$\text{ط ب} = \frac{٢}{١٢} (\text{أ ب}) = \frac{٢}{١٢} (٩) = ٢,٢٥$$

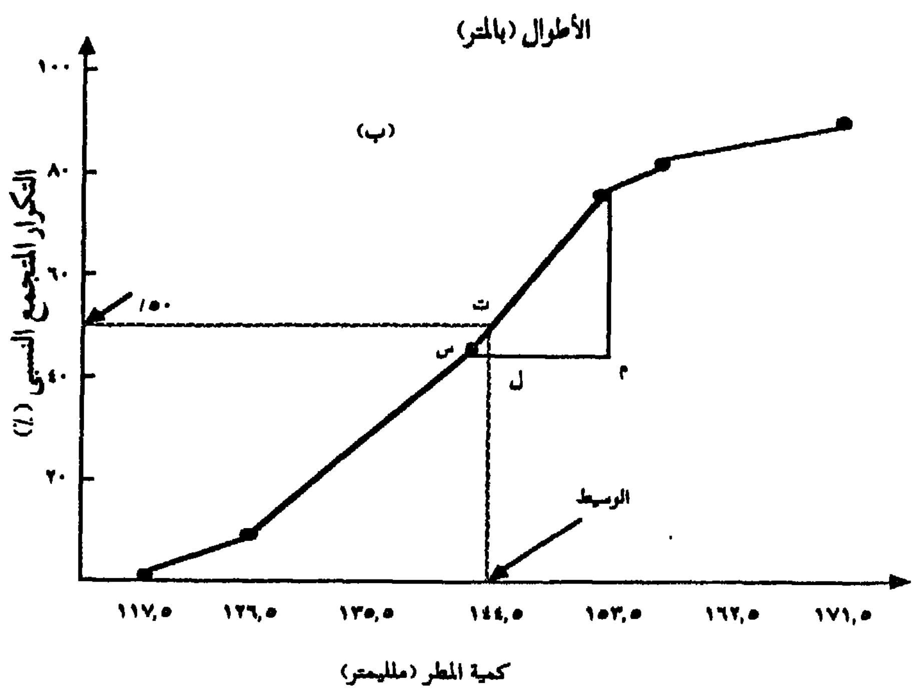
وتكون قيمة الوسيط (طـ) = قيمة ب على الاحداثى السينى + قيمة ب طـ

= ١٤٤,٥ + ٢,٢٥ = ١٤٦,٧٥

= ١٤٦,٨ ملليميتراً (إلى أقرى نسبة من عشرة من المليمتر)

ويمكن قراءة هذه القيمة بشكل تقريبى مباشرة من الرسم.





شكل رقم (٤-١): تعين الوسيط بيانياً من المدرج التكرارى (أ)
والمنحنى المتجمع النسبى الصاعد

أما إذا أردنا تعيين الوسيط من المنحى التكرارى المتجمع النسبى الصاعد المقابل لكميات الأمطار فى المثال السابق فإننا نقوم برسم محورين أحدهما رأسى يمثل التكرارات المتجمعة الصاعدة وآخر أفقى يمثل كميات الأمطار (الحدود الحقيقية للفئات) كما فى شكل رقم (٤ - ١ب). وبما أن الوسيط هو القيمة التى يسبقها ويليها عدد متساو من التكرارات، فإننا نعين على المحور الرأسى القيمة التى تمثل نصف التكرارات الصاعدة (٥٠٪). ثم نمد خطاً أفقياً منها يتقاطع مع المنحنى التكرارى المتجمع فى نقطة (ت) التى يمثل احداثيها السينى قيمة الوسيط. وللحصول على هذه القيمة فإننا نلاحظ من المثلثين المتماثلين ت ل س، ص م س أن:

$$\frac{\text{س ل}}{\text{س م}} = \frac{\text{ت ل}}{\text{ص م}}$$

وبما أن س م = طول الفئة = ٩

ت ل = (الاحداثى الصادى ت - الاحداثى الصادى ل)

ص ل = (الاحداثى الصادى ص - الاحداثى الصادى م)

فإن:

$$\frac{\text{س ل}}{٩} = \frac{٪٥٠ - ٪٤٢,٥}{٪٧٢,٥ - ٪٤٢,٥} = \frac{٧,٥}{٣٠} = \frac{١}{٤}$$

$$\therefore \text{س ل} = \frac{٩}{٤} = ٢,٢٥$$

وبهذا فإن الوسيط = ١٤٤,٥ + س ل

= ١٤٤,٥ + ٢,٢٥

= ١٤٦,٧٥

أو أن الوسيط = ١٤٦,٨ إلى أقرب عشر المللميتر. وهذه القيمة يمكن قراءتها بالتقريب من الرسم البيانى.

شبيهات الوسيط:

هناك مقاييس شبيهة بالوسيط تشترك معه فى طريقه حسابها ولكنها ليست من المتوسطات مثل الربيع Quartile والعشير Decile والمئين Centile. وسنكتفى بمعرفة طريقة حساب كل من الربيع الأدنى والربيع الأعلى نظراً لاستخدامها فى حساب مقاييس التشتت فيما بعد.

الربيع الأدنى والربيع الأعلى:

يعرف الربيع الأدنى Lower quartile بأنه تلك القيمة التى تقسم مجموعة البيانات إلى قسمين بحيث يسبقها ربع المفردات ويليها ثلاثة أرباع المفردات، ويرمز له ($ر_١$). أما الربيع الأعلى Upper quartile فيعرف بأنه القيمة التى تقسم البيانات إلى قسمين أيضاً بحيث تسبقها ثلاثة أرباع المفردات ويليها ربع المفردات، ويرمز له ($ر_٣$). ويمكن إيجاد ترتيب وقيمة كل من الربيع الأدنى والأعلى للبيانات غير المبوبة والمبوبة باتباع نفس الطريقة التى استخدمناها لإيجاد الوسيط، كما يمكن إيجادهما بيانياً من رسم المنحنى التكرارى المتجمع الصاعد أو الهابط للبيانات. وستوجد الربيعين الأدنى والأعلى لكميات الأمطار الساقطة على ٤٠ مرصد جوى (جدول رقم ٤-٨) كمايلى:

١- نوجد أولاً ترتيب الربيع الأدنى وهو = $\frac{\text{مجموع التكرارات} \times ١}{٤}$

$$= \frac{٤٠ \times ١}{٤} = ١٠$$

٢- نحدد الفئة التى يقع الربيع الأدنى من جدول رقم: (٤-٨) ونجدها الفئة (١٤٤ فأقل).

٣- نوجد قيمة الربيع الأدنى باستخدام الصيغة الآتية:

الربيع الأدنى = الحد الأدنى لفئة الربيع الأدنى +

$$\left[\left(\frac{\text{ترتيب الربيع الأدنى} - \text{التكرار المتجمع الصاعد السابق لفئة الربيع الأدنى}}{\text{التكرار الأصلى لفئة الربيع الأدنى}} \right) \times \text{طول الفئة} \right]$$

وبهذا فإن:

$$\text{الربيع الأدنى} = ١٣٦,٥ + \left(\frac{١٠ - ٨}{٩} \right) \times ٩$$

$$= ١٣٦,٥ + ٢ = ١٣٨,٥ \text{ ملليمتراً}$$

ومعنى هذا أن هناك عشرة مراصد تسقط عليها أمطار كميتها ١٣٨,٥ ملليمتراً أو أقل . وبنفس الطريقة يمكن إيجاد الربيع الأعلى كمايلى:

١ – نوجد ترتيب الربيع الأعلى وهو = $\frac{\text{مجموع التكرارات} \times ٣}{٤}$

$$= \frac{٤٠ \times ٣}{٤} = ٣٠$$

٢ – نحدد فئة الربيع الأعلى من جدول التوزيع التكرارى (جدول رقم ٤–٨) فنجدها الفئة (١٥٤ فأقل)

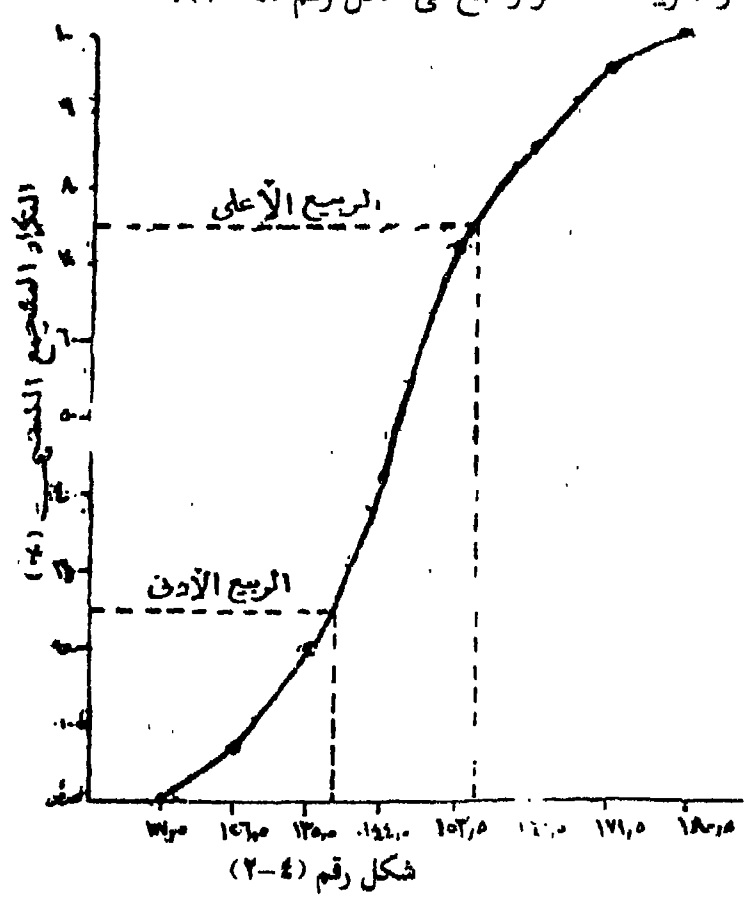
٣ – نوجد قيمة الربيع الأعلى باستخدام نفس الصيغة السابقة لإيجاد الربيع الأدنى.

$$\text{الربيع الأعلى} = ١٥٤,٥ + \left(\frac{٣٠ - ٢٩}{٥} \right) \times ٩$$

$$= ١٥٤,٥ + ١,٨$$

$$= ١٥٦,٣ \text{ ملليمتراً}$$

ولإيجاد الربيعين الأدنى والأعلى بيانياً نتبع نفس الخطوات التى انبعت لإيجاد الوسيط من منحنى التكرار المتجمع الصاعد أو الهابط فمن الشكل رقم (٤-٢) يمكن تحديد قيمة الربيع الأدنى بأن نعين ترتيبه على المحور الرأسي ونرسم خطاً أفقياً ليقابل المنحنى فى نقطة يمثل احداثيها السينى القيمة المطلوبة وذلك عن طريق أسقاط عمود من هذه النقطة ليقابل المحور الأفقى عند قيمة ١٣٨,٥ ملليمتر تقريباً. وبالنسبة لإيجاد قيمة الربيع الأعلى فإننا نمين ترتيبه على المحور الرأسى ونتبع نفس الطريقة لتحديد قيمته على المحور الأفقى فنجد أن قيمته تساوى ١٥٦,٥ ملليمتر تقريباً كما هو واضح فى شكل رقم (٤- ٢):



شكل رقم (٤-٢)

تعيين الربيع الأدنى والأعلى بيانياً من المنحنى المتجمع النسبى الصاعد

ثالثاً: المنوال Mode

المنوال، أو مايعرف أحياناً بالشائع، وهو أحد مقاييس النزعة المركزية التى تعطى تصوراً عاماً لتركز قيم مجموعة من المفردات حول قيمة متوسطة تمثل اتجاها لهذه المجموعة كما تعتبر أحد خصائصها ومعالمها. ويمكن أن يعرف المنوال بأنه القيمة التى تحدث أكثر من غيرها فى مجموعة من البيانات أى هو القيمة الأكثر تكرار أو شيوعاً فى التوزيع. فإذا كانت لدينا البيانات التالية للأحجام السكانية (بالألف نسمة) لعدد من المدن فى اقليم ما:

٧، ١٥، ٦، ٢٥، ٥،٧، ٢٦، ٧، ١٥، ١٦

نجد أن القيمة (٧) تكررت أكثر من غيرها ولذلك فإن المنوال لهذه المجموعة يكون مساوياً للمدينة التى حجم سكانها ٧ ألاف نسمة. ويطلق على مثل هذا التوزيع الذى به قيمة واحدة هى الأكثر تكراراً إسم التوزيع أحادى المنوال Unimo-dal أما إذا كانت هناك قيمتان متساويتان فى تكراراتهما فإن التوزيع يسمى بالتوزيع مزدوج (ثنائى) المنوال Bimodal كما هو واضح فى البيانات التالية للأحجام السكانية (بالألف نسمة) لعدد من المدن فى اقليم آخر:

٤، ٤، ٥، ٥، ٥، ٦،٥، ٧، ٩، ٩، ٩، ٩، ١٠، ١١

يلاحظ أن كل من القيم ٥، ٩ هما القيمتان الأكثر تكرار من غيرها من القيم الأخرى أى أن البيانات السابقة منوالين يتملثان فى المدينتين اللتين حجم سكانها ٥، ٩ آلاف نسمة. وقد لا يوجد قيمة منوالية فى مجموعة البيانات عندما لاتتكرر قيمة ما أكثر من غيرها كما يلاحظ من بيانات الحجم السكانى (بالألف نسمة) لعدد من المدن فى أقليم ثالث.

٨، ٤، ١٣، ٩، ٧، ١٢، ٥، ١٠

إيجاد المنوال للبيانات المبوبة:

تختلف قيمة المنوال المحسوبة من التوزيع التكرارى عن قيمته المحسوبة من البيانات غير المبوبة. ويحسب المنوال من البيانات المبوبة فى جداول التوزيعات

التكرارية وذلك بتحديد الفئة التى تضم أكبر عدد من التكرارات وتعرف بالفئة المنوالية Modal Class ويمكن اعتبار مركز هذه الفئة منوالا للتوزيع فى حالة تساوى التكرارات المناظرة للفئتين قبل وبعد الفئة المنوالية.

وهناك عدة طرق لتحديد قيمة المنوال داخل الفئة المنوالية. وتعتمد الطريقة الأولى فى حساب المنوال على معرفة تكرار كل من الفئتين المحيطين بالفئة المنوالية، وبالاقتباس من طريقة الرافعة فإن الفئة المنوالية تمثل الرافعة ويمثل تكرار الفئة قبل المنوالية القوة وتكرار الفئة بعد المنوالية المقاومة. وعلى هذا الأساس يتحدد موضع (قيمة) المنوال عند نقطة إرتكاز هذه الرافعة كما فى الشكل رقم (٤-٣) اعتماداً على المثال التالى:

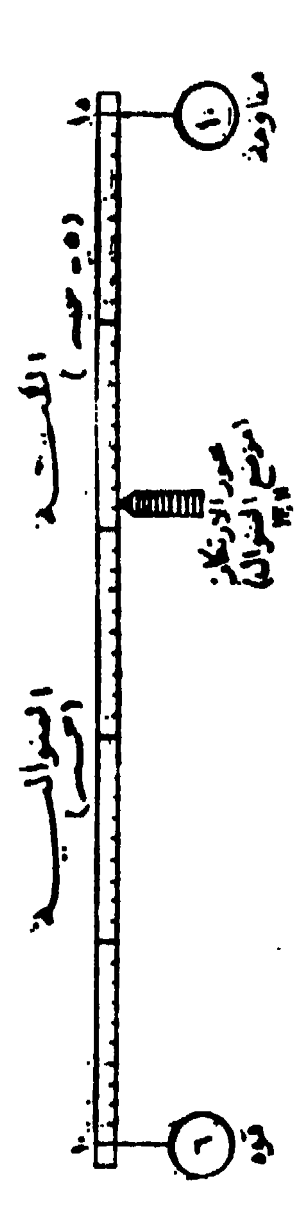
جدول رقم (٤-٩)

حساب المنوال لمساحة الأحواض

الزراعية المزروعة قطنا فى أحد المحافظات «بالفدان»

الفئات (المساحة بالفدان)	التكرار (عدد الأحواض)	مراكز الفئات
١-	٤	٢,٥
٥-	٦	٧,٥
١٠-	١٨	١٢,٥ (الفئة المنوالية)
١٥-	١٠	١٧,٥
٢٠-	٨	٢٢,٥
٢٥-	٢	٢٧,٥

فالأحواض الزراعية التى مساحتها من ١٠ لأقل من ١٥ فداناً هى الفئة المنوالية حيث أنها الأكثر تكراراً (١٨ تكراراً) والمنوال لهذا التوزيع هو ١٢,٥ حيث أنها القيمة المركزية للفئة ١٠ لأقل من ١٥. ولتحديد موضع (قيمة) المنوال داخل هذه الفئة يطبق قانون الرافعة.



شكل رقم (٤-٣)
تحديد موضع المنوال داخل الفئة المنوالية
«بطريقة الرافعة»

من الشكل السابق نرى أن المنوال يبعد مسافة قدرها س عن بداية الفئة (١٠) وبالتالى سيبعد مسافة قدرها (طول الفئة -س) أى (٥-س) عن نهاية الفئة (١٥)، ومن قانون الرافعة نجد أن:

القوة × ذراعها = المقاومة × ذراعها

التكرار السابق للفئة المنوالية × س = التكرار اللاحق للفئة المتوالية × (٥ -س)

٦ × س = ١٠ × (٥ - س)

∴ ٦س = ٥٠ - ١٠ س

٦ = ٥٠

$$\therefore \quad س = \frac{٥٠}{١٦} = ٣,١١$$

وعلى ذلك فإن:

المنوال = الحد الأدنى للفئة المنوالية + قيمة محور الارتكاز (س)

= ١٠ + ٣,١١ = ١٣,١١ فداناً

أما الطريقة الثانية لحساب المنوال من جداول التوزيعات التكرارية فقد اقترحها بيرسون K. Pearson وتعرف بطريقة الفروق (بيرسون). وتقوم هذه الطريقة على أساس أن الذى يحدد موضع (قيمة) المنوال داخل الفئة المنوالية هو الفرق بين تكرار الفئة المنوالية وتكرارى الفئتين السبقة واللاحقة لها. وعلى ذلك يتحدد موضع المنوال بنسبة الفرق بين التكرارات على طرفى الفئة المنوالية وباستخدام العلاقة.

$$المنوال = ف_م + \left[\frac{ك - ك_١}{(ك - ك_١) + (ك - ك_٢)} \right] \times ل \quad ..... (٤-١٢)$$

حيث أن:

$ف_م$ = الحد الأدنى للفئة المنوالية.

ك = تكرار الفئة المنوالية.

$ك_{١}$ = تكرار الفئة السابقة لفئة المنوال.

$ك_{٢}$ = تكرار الفئة اللاحقة لفئة المنوال.

ل = طول الفئة.

ولإيجاد المنوال بهذه الطريقة من الجدول رقم (٤-٩) نجد أن :

الفرق بين تكرار الفئة المنوالية والفئة السابقة لها ($ك - ك_{١}$) = ١٨ - ٦ = ١٢

الفرق بين تكرار الفئة المنوالية والفئة اللاحقة لها ($ك - ك_{٢}$) = ١٨ - ١٠ = ٨

ويحسب المنوال على أساس أنه القيمة التى تقسم الفئة المنوالية (١٠ إلى أقل من ١٥) بنسبة: ١٢ :٨. فإذا كانت المسافة التى يبعد بها المنوال عن بداية الفئة هى س فإن المسافة التى يبعد بها عن نهاية الفئة هى (طول الفئة - س) أى (٥-س). وبهذا فإن:

س: (٥ - س) يجب أن تكون كنسبة ١٢ :٨ أى أن:

$$\frac{\text{س}}{(\text{٥} - \text{س})} = \frac{\text{١٢}}{\text{٨}}$$

∴ ٨س = ١٢ (٥ - س)

٨س = ٦٠ - ١٢س

∴ ٢٠س = ٦٠

∴ س = ٣

والمنوال فى هذه الحالة = ١٠ + ٣ = ١٣ فداناً.

ويمكن الحصول على هذه النتيجة مباشرة باستخدام الصيغة (٤-١٢):

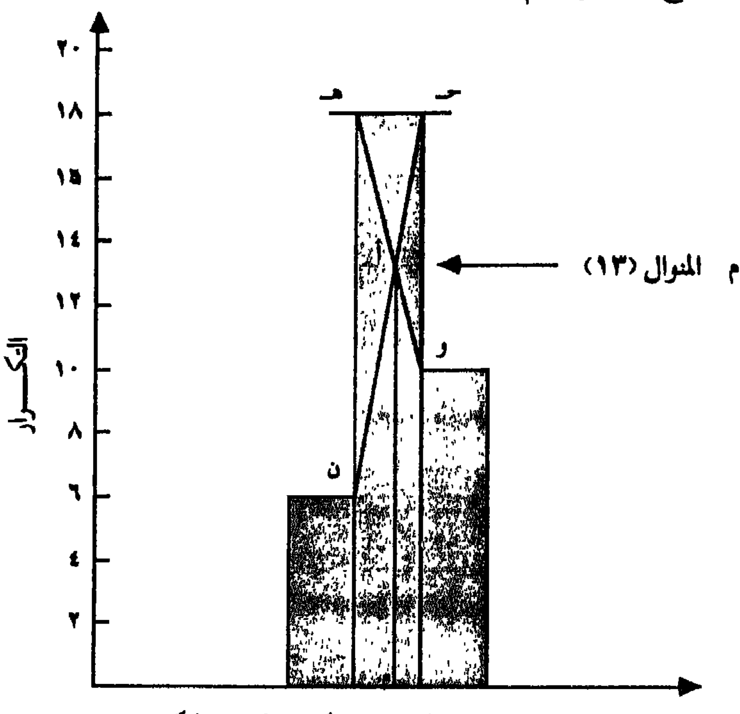
$$\text{المنوال} = \text{١٠} + \left(\frac{\text{١٢}}{\text{١٢} + \text{٨}} \right) \times \text{٥} = \text{١٠} + \frac{\text{٦٠}}{\text{٢٠}}$$

= ١٠ + ٣ = ١٣ فداناً

إيجاد المنوال بيانياً:

المنوال مثل الوسيط يمكن الحصول عليه من الرسم البيانى للمدرج التكرارى بتحديد الثلاثة مستطيلات التى يمثل المستطيل الأوسط منها تكرار الفئة المنوالية ويمثل كل من المستطلين الآخرين تكرار الفئة السابقة والفئة اللاحقة لها على حدة (شكل رقم ٤-٤).

ويعين المنوال عن طريق توصيل طرفى المستطيل الذى يمثل تكرار الفئة المنوالية بطرفى المستطيلين اللذين يمثلان تكرارى الفئتين السابقة واللاحقة لها بمستقيمين، (حـ د)و (هـ و)، أى بتوصيل الطرف الأيمن العلوى لمستطيل الفئة المنوالية بالطرف الأيمن العلوى للفئة السابقة لها (حـ د)، والطرف الأيسر العلوى بالطرف الأيسر العلوى للفئة اللاحقة لها (هـ و)، والاحداثى السينى لنقطة التقاطع (أ) يساوى المنوال (م) الذى يمكن قراءة قيمته بشكل تقريبى مباشرة من الرسم كما هو موضح الشكل رقم (٤-٤).



شكل رقم (٤-٤ : طريقة تعيين المنوال بيانياً من المدرج التكرارى

العلاقة بين مقاييس النزعة المركزية:

سبق أن ذكرنا أن مقاييس النزعة المركزية الثلاثة: المتوسط الحسابى، الوسيط والمنوال هي أدوات القياس الكمى الرئيسية التى تعطى تصوراً عاماً لدرجة تركز قيم المفردات حول قيمة متوسطة معينة وتعتبر هذه القيمة هى اتجاه عام لمجموعة مفردات البيانات وأحد معالمها وخصائصها. ومن التعريف السابق أمكن معرفة كل مقياس من هذه المقاييس، كما لوحظ أن هناك علاقة تربط بين كل منها. ولهذا فمن المفيد توضيح العلاقة هذه المتوسطات فى ضوء الأمثلة الآتية:

فإذا أردنا مثلاً معرفة متوسط المساحة الزراعية المملوكة لعينة من عشرة زراع يمتلك كل منهم المساحة التالية (بالفدان):

١ ، ٢ ، ٣ ، ٣ ، ٣ ، ٣ ، ٤ ، ٤ ، ٥

وتم تمثيل هذا التوزيع بيانياً فى شكل مدرج تكرارى وممهداً عليه منحنى توزيع تكرارى ليوافق هذا التوزيع (شكل رقم ٤–٥أ) فإننا نرى أن قيم المقاييس الثلاثة تتلاقى (تتطابق) فى قيمة واحدة وهى الرقم (٣)، حيث أن:

المتوسط الحسابى = $\bar{س}$ = $\frac{مجـ\ س}{ن}$

= $\frac{٣٠}{١٠}$ = ٣ أفدنة

الوسيط = $\frac{٣ + ٣}{٢}$ = ٣ أفدنة

المنـوال = ٣ أفدنة (القيمة الأكثر تكراراً أو شيوعاً)

ويلاحظ من الشكل أيضاً أن منحنى التوزيع التكرارى يعد متزناً تماماً على كلا جانبى قيم المتوسطات، ويعرف فى هذه الحالة بمنحنى التوزيع المتماثل الذى يعد، بالإضافة إلى تطابق قيم المقاييس الثلاثة عليه، سمة مميزة لما يعرف بالتوزيع التكرارى المعتدل (الطبيعى) Normal Frequency Distribution الذى تعتمد عليه

كثير من الأساليب الاحصائية على الرغم من ندرة حدوثه من الوجهة العملية. إذ كثيراً ماتجد أن قيم البيانات لا تشكل توزيعاً معتدلاً (متزنا) حول قيم المتوسط، بل تمثل توزيعاً غير معتدل Non- Normal Distribution مثل البيانات التالية التى تبين المساحات الزراعية (بالفدان) المملوكة لعينة من عشرة زراع:

١، ١، ٢، ٢، ٢، ٣، ٤، ٤، ٥

فإننا نجد أن المتوسطات الثلاثة لهذه البيانات تختلف فى قيمتها وذلك على النحو التالى:

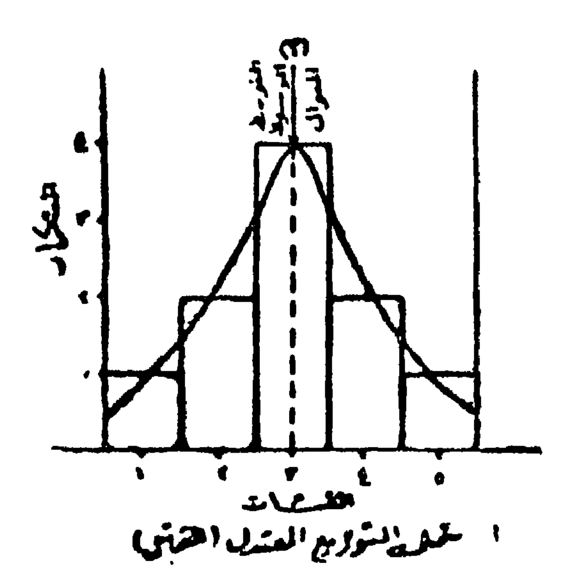
$$\text{المتوسط الحسابى} = \bar{س} = \frac{\text{مجـ س}}{ن} = \frac{٢٧}{١٠} = ٢,٧ \text{ فدانا}$$

أما استخراج الوسيط والمنوال فيكون عن طريق جدولة القيم أى إعادة ترتيبها أو تبويبها على النحو التالى:

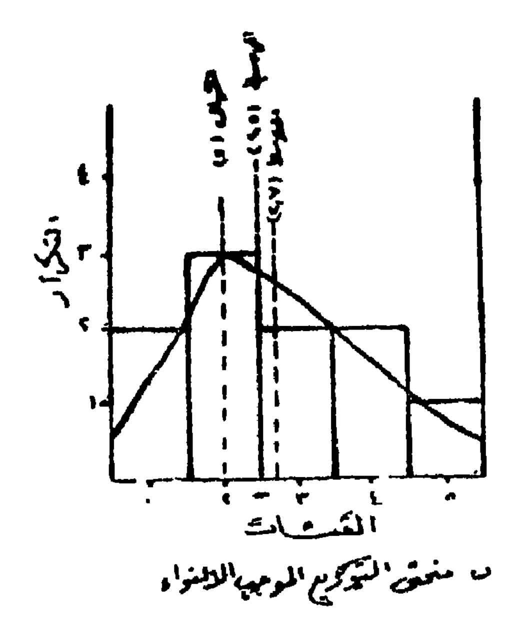
القيمة	عدد مرات حدوثها (التكرار)
١	٢
٢	٣
٣	٢
٤	٢
٥	١

فالوسيط كما ذكرنا هى القيمة الوسطى. وبما أن عدد القيم يساوى ١٠، فإن الوسيط هو متوسط القيمتين الوسطيتين:

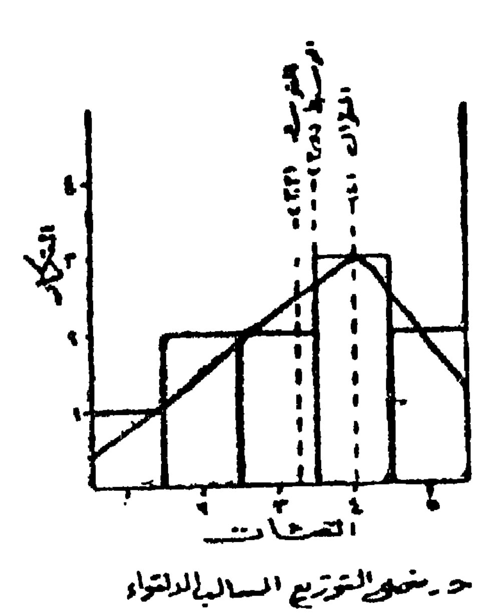
$$\text{الوسيط} = \frac{٢ + ٣}{٢} = ٢,٥ \text{ فدانا}$$



أ منحنى التوزيع المعتدل (المتماثل)



ب منحنى التوزيع الموجب الالتواء



ج منحنى التوزيع السالب الالتواء

شكل رقم (٤-٥)

مقاييس النزعة المركزية وعلاقتها بأنواع التوزيعات التكرارية

أما المنوال فهو القيمة التى تكررت أكثر من غيرها وهى القيمة (٢ فدان) التى تكرر حدوثها ثلاث مرات.

وقد تم توضيح القيم الثلاث السابقة على المدرج التكرارى والمنحنى التكرارى للتوزيع فى الشكل رقم (٤-٥ب) ويلاحظ من هذا الشكل أن المقاييس الثلاثة تبدو مختلفة عن بعضها ولا تتفق فى قيمة واحدة، كما رأينا فى المثال السبق، إذ يعد المتوسط الحسابى (٢,٧) أكبر قيمة بينما المنوال أصغرها (٢). كما يلاحظ أيضاً أن منحنى التوزيع غير متوازن (أى ملتو)- إذ أن له قمة تميل إلى الجانب الأيسر من المركز، وذيلاً يتجه صوب اليمين، ويسمى هذا بالالتواء الموجب Positive Skewness. وتتميز التوزيعات مثل هذا النوع بأن النمط النسبى لتوزيع المتوسطات الثلاثة على منحنياتها يكون على نحو أن الوسيط يقع فى منتصف التوزيع بينما يقع المنوال على يساره والمتوسط الحسابى على يمينه كما يبدو من الشكل رقم (٤- ٥ب).

كذلك تختلف المتوسطات الثلاثة فى قيمتها إذا كانت قيم البيانات تمثل توزيعاً غير معتدل وبصورة عكسية لبيانات المثال السابق. فمثلاً إذا كانت لدينا أيضاً بيانات بالمساحات الزراعية (بالفدان) المملوكة لعينة مكونة من عشرة من الزراع على النحو التالى:

١، ٢، ٣، ٣، ٤، ٤، ٤، ٥، ٥، ٥

فإن:

$$\text{المتوسط الحسابى} = \bar{س} = \frac{\text{مجـ س}}{ن} = \frac{٣٣}{١٠} = ٣,٣ \text{ فدانا}$$

أما بالنسبة لحساب لكل من الوسيط والمنوال فإن للقيم يعاد ترتيبها كمايلى:

القيمة	عدد مرات حدوثها (التكرار)
١	١
٢	٢
٣	٢
٤	٣
٥	٢

$$\text{الوسيط} = \frac{٣ + ٤}{٢} = ٢,٥ \text{ فدانا}$$

المنوال = ٤ فداناً (القيمة الأكثر تكراراً أو شيوعاً)

وهكذا تبدو أيضاً المتوسطات مختلفة عن بعضها البعض، فيعد المنوال أكبرها والمتوسط الحسابي أصغرها، وهو عكس ما رأينا في التوزيع السابق. ومن الشكل رقم (٤– ٥حـ) الذى يوضح هذه القيم على مدرج تكرارى ممهداً عليه منحنى تكرارى للتوزيع، يبدو الأخير غير متوازن حيث يكون مائلاً نحو اليسار أن أى قمته تتجه إلى يمين المركز، وذيله صوب اليسار. وفى هذه الحالة يعرف بالتوزيع السالب الالتواء Negative Skewness. والتوزيعات الملتوية السالبة خاصية مميزة من حيث النمط النسبى لتوزيع المتوسطات الثلاثة على منحنياتها حيث نجد أن الوسيط يقع فى منتصف التوزيع والمتوسط الحسابى على يساره والمنوال على يمينه.

ومما تجدر الإشارة إليه هنا، أنه من الشكل رقم (٤–٥) يمكن استنتاج علاقة تقريبة عامة بين مقاييس النزعة المركزية الثلاثة للتوزيعات غير المتماثلة ، وهذه العلاقة يمكن التعبير عنها حسابياً على النحو التالى:

أولاً: التوزيعات الملتوية الموجبة:

المنوال = المتوسط الحسابى – ٣ (المتوسط الحسابى – الوسيط)

وفى تطبيق هذه العلاقة على المثال الموضح فى شكل رقم (٤–٥ب) نجد أن:

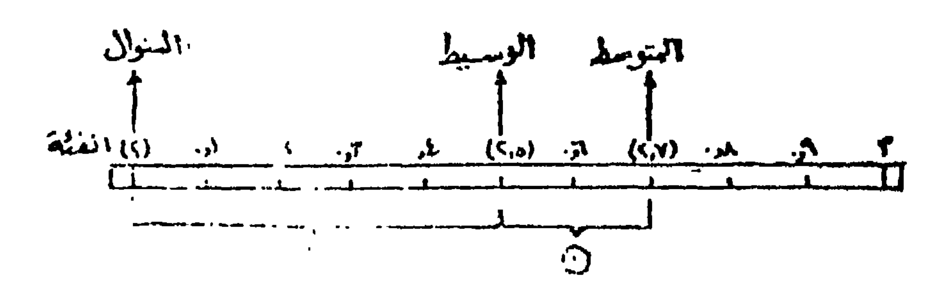
المنوال = ٢,٧ − ٣ (٢,٧ − ٢,٥)

= ٢,٧ − (٣ × ٠,٢)

= ٢,٧ − ٠,٦

= ٢,١ فداناً

وقد سبق أن عرفنا أن القيمة المنوالية من التوزيع هى (٢)، والفرق بين القيمتين بسيط جداً، كما أن هذا يعنى أن الوسيط يقع إلى الخلف من المتوسط الحسابى بمقدار الثلث تقريباً وأمام المنوال بمقدار الثلثين تقريباً كما هو واضح فى الشكل رقم (٤−٦).



شكل رقم (٤−٦)

العلاقة بين النزعة المركزية الثلاثة فى التوزيعات التكرارية الموجبة الالتواء

جدول رقم (٤-١٠)

كمية الأمطار السنوية (بالبوصة) فى مرصد

Bidston, Birkenhead - انجلترا فى الفترة ١٩٠١ - ١٩٣٠

السنة	كمية الأمطار بالبوصة	السنة	كمية الأمطار بالبوصة	السنة	كمية الأمطار بالبوصة
١٩٠١	٢٥,٩١	١٩١١	٢٥,٢٧	١٩٢١	٢٢,٤٧
٢	٢٥,٥٧	١٢	٣٠,١٧	٢٢	٢٥,٩٧
٣	٣٤,٤٢	١٣	٢٥,٧٨	٢٣	٣٠,٩٢
٤	٢٥,٨١	١٤	٢٦,٠٢	٢٤	٣٢,٨٧
٥	٢٤,٠١	١٥	٢٦,٨٣	٢٥	٢٨,٠٠
٦	٢٨,٠٨	١٦	٢٤,٨٧	٢٦	٢٨,٩٥
٧	٢٦,٥٧	١٧	٣٠,٥٩	٢٧	٣٤,٨١
٨	٢٨,٩٠	١٨	٣١,٩٣	٢٨	٢٩,١١
٩	٢٨,٣٥	١٩	٢٩,١٢	٢٩	٢٥,١٥
١٠	٢٨,٥٩	٢٠	٣٣,٣٤	٣٠	٣٦,٥٠

يبدو من بيانات الجدول السابق أن كمية الأمطار السنوية تتراوح من ٢٤,٢ بوصة و ٣٦,٥٠ بوصة، كما تمثل بيانات هذا الجدول متغيراً مستمراً (متصلاً) Continuous Variate. وبحساب المتوسط الحسابى من القيم فى الجدول وجد أن:

$$\bar{س} = \frac{مجـ\ س}{ن} = \frac{٨٥٣,٦٣}{٣٠} = ٢٨,٤٥ \text{ بوصة}$$

ولحساب كل من الوسيط والمنوال فقد أعيد ترتيب البيانات فى الجدول السابق فى منظومة Array أى ترتيبها ترتيباً تصاعدياً حسب مقدارها، كما تم تبويبها أيضاً على شكل توزيع تكرارى بعد أن تم تجميع القيم فى فئات مناسبة وعدد مرات

حدوث كل قيمة (التكرارات) أمام كل فئة. ولما كانت البيانات تمثل متغيراً متصلاً فإن كل القيم تشتمل على كسور عشرية، ومن ثم فلابد أن تصمم حدود الفئات لتعطى متغيراً مستمراً فلا تكون مثلاً (٢١ -٢٢)، (٢٣- ٢٤).. الخ، بل يجب أن تكون (٢١ إلى ٢٢,٩٩)، (٢٣ - ٢٤,٩٩) وهكذا. ويبين الجدول رقم (٤-١١) والشكل رقم (٤-٧) هذه البيانات.

جدول رقم (٤-١١)

كمية الأمطار السنوية (بالبوصة) فى مرصد

Bidston, Birkenhead- انجلترا فى الفترة ١٩٠١ - ١٩٣٠

مرتبة بحسب مقدار القيمة

كمية الأمطار مرتبة حسب قيمتها	تحويل القيم إلى فئات	عدد مرات الحدوث التكرار
٢٢,٤٧	٢١ - ٢٢,٩٩	١
٢٤,٠١	٢٣ - ٢٤,٩٩	٢
٢٤,٨٧	٢٥ - ٢٦,٩٩	١٠
٢٥,١٥	٢٧ - ٢٨,٩٩	٦
٢٥,١٨	٢٩ - ٣٠,٩٩	٥
٢٥,١٩	٣١ - ٣٢,٩٩	٢
٢٥,٢٧	٢٣ - ٣٤,٩٩	٣
٢٥,٥٧	٣٥ - ٣٦,٩٩	١
٢٥,٨٧	الفئة المنوالية ٢٥ - ٢٦,٩٩	
٢٥,٩٧		
٢٦,٠٢		
٢٦,٥٧		
٢٦,٨٣		
٢٨,٠٠		

ثانياً: التوزيعات الملتوية السالبة:

يمكن أن يعبر عن العلاقة التقريبية العامة التي تربط مقاييس النزعة المركزية الثلاثة (المتوسط الحسابي، الوسيط، المنوال) للتوزيعات السالبة الالتواء بصورة عكسية لنفس العلاقة مع التوزيعات الملتوية الموجبة وذلك على النحو التالي:

المنوال = المتوسط الحسابي + ٣ (الوسيط − المتوسط الحسابي)

وبتطبيق هذه المعادلة على بيانات الشكل رقم (٤–٥ جـ) نجد أن.

المنوال = ٣,٣ + ٣ (٣,٥ − ٣,٣)

= ٣,٣ + (٣ × ٠,٢)

= ٣,٣ + ٠,٦

= ٣,٩ فداناً

وسبق القول بأن القيمة المتوالية من للتوزيع هى القيمة (٤). والفرق بين القيمتين كما هو واضح بسيط جداً.

وهناك أيضاً علاقة تقريبية عامة بين المتوسطات الثلاثة يمكن بواسطتها حساب المتوسط الحسابى للتوزيعات غير المتماثلة (الموجبة أو السالبة). وهذه العلاقة يعبر عنها كمايلى:

$$\text{المتوسط الحسابى} = \frac{\text{٣ (الوسيط)} - \text{(المنوال)}}{\text{٢}}$$

وبتطبيق هذه العلاقة على بيانات كل من الشكلين رقم (٤– ٥ب)، (٤–٥جـ) نجد أن:

$$\text{١− المتوسط الحسابى} = \frac{(\text{٣} \times \text{٢,٥}) - (\text{٢})}{\text{٢}}$$

$$= \frac{\text{٥,٥}}{\text{٢}} = \text{٢,٧٥ فدانا}$$

٢- المتوسط الحسابى = $\frac{(٣ \times ٣,٥) - (٤)}{٢}$

= $\frac{٦,٥}{٢}$ = ٣,٢٥ فدانا .

وكما نرى فإن القيمتين ٢,٧٥ و ٣,٢٥ قريبتان جداً من مثيلتهما فى التوزيع ٢,٧ ، ٣ ، ٣ على الترتيب مما يدل على دقة هذه العلاقة بين المتوسطات الثلاثة.

أمثلة تطبيقة:

بعد ذلك العرض العام الذى أوضحنا فيه العلاقة بين مقاييس النزعة المركزية (المتوسط الحسابى، الوسيط والمنوال)، نحاول الآن ابراز هذه العلاقة بصورة حية وذلك بأخذ بيانات نوعية فى مجال الدراسات الجغرافية (الطبيعية والبشرية) وتطبيق الطرق السبق ذكرها لحساب هذه المقاييس الثلاثة عليها. فالجدول رقم (٤-١٠) يبين كمية الأمطار السنوية (بالبوصة) التى سجلها مرصد Bidston, Birkenhead - انجلترا وذلك لمدة ثلاثين عاماً من ١٩٠١ إلى ١٩٣٠. والمطلوب تمثيل البيانات فى الجدول بيانياً، وإيجاد المتوسطات الثلاثة حسابياً وتوضيح مابينها من علاقة.

٢٨,٠٨	
	← الوسيط (٢٨,٢٧)
٢٨,٤٥	
٢٨,٥٩	
٢٨,٩٠	
٢٨,٩٥	
٢٩,١١	
٢٩,١٢	
٣٠,١٧	
٣٠,٥٩	
٣٠,٩٢	
٣١,٩٣	
٣٢,٨٧	
٣٣,٣٤	
٣٣,٤٢	
٣٤,٨١	
٣٦,٥٠	
٨٥٣,٦٣	المجموع =

ولما كان عدد قيم المفردات فى الجدول السابق ٣٠ قيمة فإن الوسيط سيتحدد فيما بين القيمتين الخامسة عشر والسادسة عشر، أى فى منتصف المسافة بين ٢٨,٨ بوصة و ٢٨,٤٥ بوصة أى أن:

$$\text{الوسيط} = \frac{٢٨,٤٥ - ٢٨,٠٨}{٢}$$

وللمقارنة يمكن حساب الوسيط من التوزيع التكرارى للبيانات باستخدام الصيغة (٤-١١) فنجد أن:

$$\text{الوسيط} = ٢٧,٠٠ +$$

$$= ٢٧,٠٠ + (٢ \times ٠,٣٣)$$

$$= ٢٧,٠٠ + ٠,٦٦$$

$$= ٢٧,٦٦$$

وكما نرى تختلف القيمة الأخيرة عن القيمة الدقيقة السابقة للوسيط. ويرجع هذا الخطأ – الذى يعتبر أكبر من الخطأ المعتاد فى مثل هذه الحالة – إلى وقوع كل التكرارات الأصلية الستة فى الفئة الوسيطية فى النصف الأعلى لهذه الفئة.

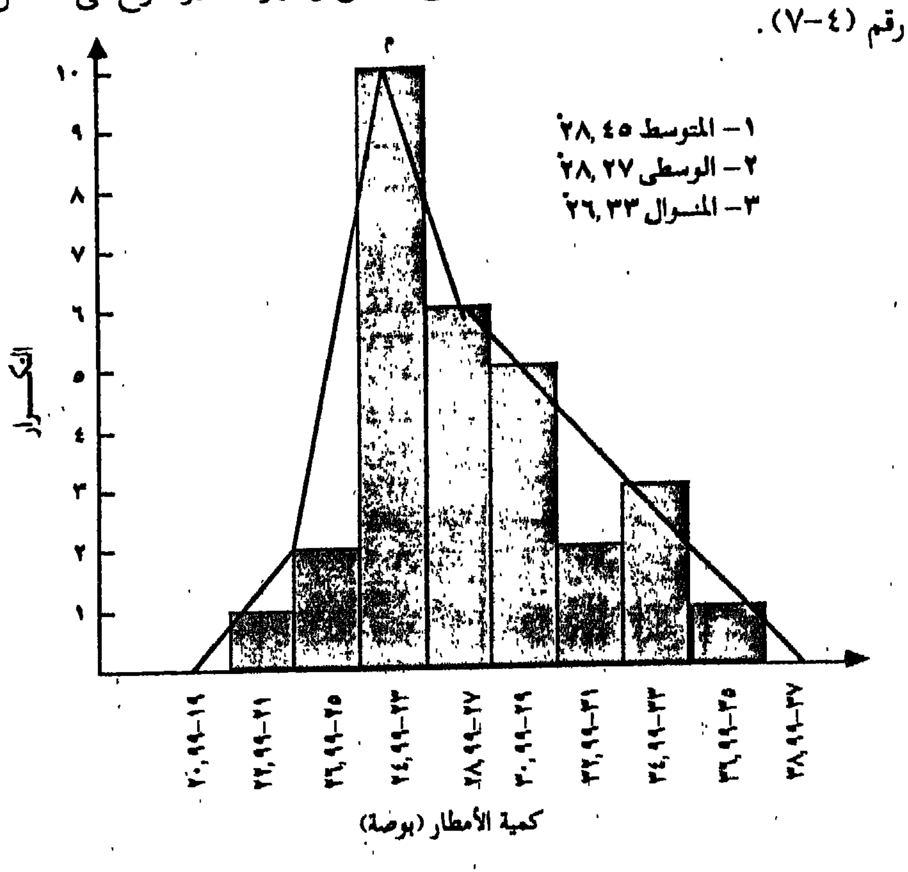
ومن الجدول رقم (٤-١١) والشكل رقم (٤-٧) يتضح أن الفئة ٢٥ – ٢٦,٩٩ هى الفئة المنوالية التى يوجد بها أكبر عدد للتكرارات (١٠). ويمكن بالتالى بعد تحديد الفئة التى يقع فيها المنوال حساب قيمته باستخدام العلاقة (٤-١٢) على النحو التالى:

$$\text{المنوال} = ٢٥,٠٠ \left[\frac{١٠ - ٢}{(١٠ - ٢) + (١٠ - ٦)} \right] \times ٢$$

$$= ٢٥,٠٠ + \frac{٨}{١٢} \times ٢$$

$$= ٢٥,٠٠ + ١,٣٣ = ٢٦,٣٣ \text{ بوصة}$$

وهكذا تبدو المتوسطات الثلاثة مختلفة عن بعضها البعض، إذ نجد أن المتوسط الحسابى هو أكبر قيمة بينما المنوال أقلها. وقد تم تمثيل هذه القيم على المدرج التكرارى والمنحنى التكرارى للتوزيع فى الشكل رقم (٤-٧) ومنه يبدو أن المنحنى ملتو التواء موجباً بسيطاً. ويستدل على ذلك، كما سبق القول من الصفة المميزة لكل التوزيعات الملتوية الموجبة وهى أن النمط النسبى للمتوسطات الثلاثة (أو توزيعها) على المنحنى التكرارى يأخذ شكلاً يكون فيه الوسيط فى منتصف التوزيع والمنوال على اليسار والمتوسط الحسابى على اليمين ويظهر هذا بوضوح فى الشكل رقم (٤-٧).



شكل رقم (٤-٧)

المدرج التكرارى ومنحنى التوزيع التكرارى لكمية الأمطار السنوية فى مرصد بيستون - انجلترا فى الفترة ١٩٠١ - ١٩٣٠

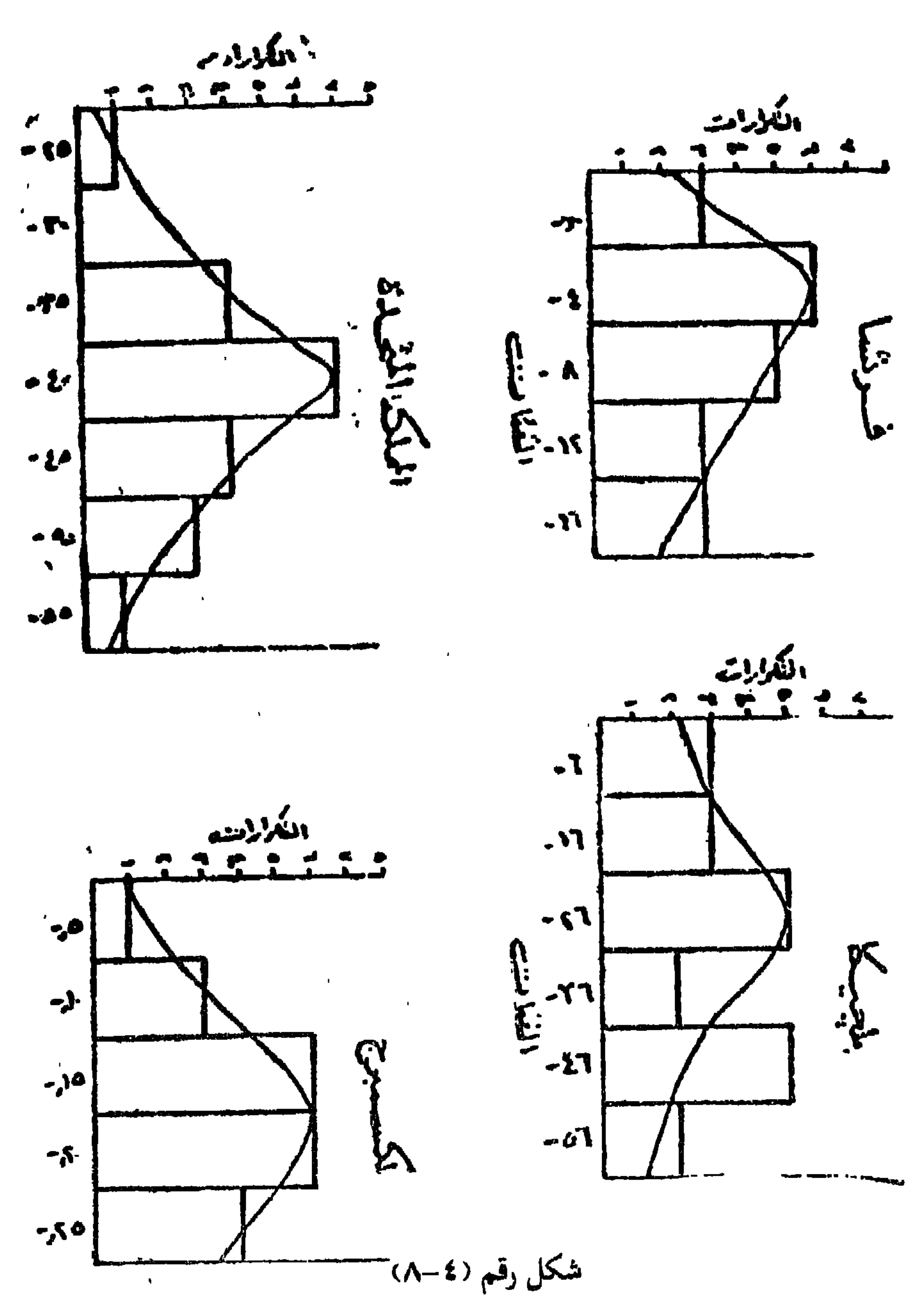
وتبدو الاختلافات فى التوزيعات التكرارية والعلاقة بين مقاييس النزعة المركزية المتوسطات الثلاثة فى بيانات تتصل بمجال الدراسة فى الجغرافية الاقتصادية، كما يظهر من بيانات الإنتاج السنوى لخام الحديد على مدى عشرين عاماً ١٩٣٨ – ١٩٥٧ فى أربع دول من غرب أوربا هى: بلجيكا، فرنسا، لكسمبرج والمملكة المتحدة فى الجدول التالى:

جدول رقم (٤-١٢)

إنتاج خام الحديد بآلاف الاطنان فى

بعض دول غرب أوربا فى الفترة من ١٩٣٨ – ١٩٥٧

السنة	بلجيكا	فرنسا	لكسمبرج	المملكة المتحدة
١٩٣٨	٦٥	١٠٢٠٣	١٥٠٦	٣٦١٥
٣٩	٦٠	١٠١٦١	١٦٣٩	٤٤١٧
٤٠	٢٩	٤١١٣	١٣٦٨	٥٤٤٩
٤١	٤٧	٣٤٦٧	١٩١٢	٥٥٢٨
٤٢	٤١	٤١٤٤	١٤٣١	٥٤٤٩
٤٣	٤٦	٥٣٥٠	١٤٧١	٥٤١١
٤٤	١٦	٢٨٦٢	٨١٦	٤٣٩٠
٤٥	١١	٢٣٤٢	٣٩٤	٤١٦٢
٤٦	١٤	٥٠٢١	٦٥٠	٣٥٧٤
٤٧	٢١	٦٠٩٩	٥٩٢	٢٩٨٤
٤٨	٣٤	٧٥٥٥	١٠٢٠	٣٩٩٠
٤٩	١٥	١٠٢٠٠	١٢٤١	٤٠٨٦
٥٠	١٦	٩٧٥٠	١١٥٤	٣٨١٢
٥١	٢٨	١١٤٤٠	١٦٨٨	٤٥٠٤
٥١	٤٧	١٣٢٣٠	٢١٧٤	٤٦١٨
٥٣	٣٥	١٣٧٩٠	٢١٥١	٤٥٠٠
٥٤	٢٩	١٤٢٤٠	١٧٦٦	٤٣٦٩
٥٥	٣٧	١٦٣٤٠	١٩٣٣	٤٤٣٧
٥٦	٥٠	١٧١٢٠	٢٠٣٤	٤٤٥٧
١٩٥٧	٤٨	١٨٧٧٠	٢٠٣٦	٤٦٣٧
المتوسط	٣٤,٤٥	٩٣١٠,٢	١٤٤٨,٨	٤٤١٨,٩٥
الوسيط	٣٤,٥	٩٩٥٥,٥	١٤٨٨,٥	٤٤٢٧,٠٠



شكل رقم (٤-٨)

المدرجات والمنحنيات التكرارية للإنتاج السنوى لخام الحديد فى كل من بلجيكا، فرنسا، لكسمبرج والمملكة المتحدة فى الفترة من ١٩٣٨ - ١٩٥٧

وكما هو واضح من الجدول السابق فقد تم حساب كل من المتوسط الحسابى والوسيط لإنتاج كل دولة بنفس الطرق السابق ذكرها مع بيانات الجدول رقم (٤-١١)، ويبدو فى الحالات الأربع أن الوسيط أكبر من المتوسط الحسابى- مما يترتب عليه ميلا بسيطاً نحو التواء سالب - بالرغم من أن التوزيع التكرارى لبيانات فرنسا لا يؤيد ذلك. ويبين الشكل رقم (٤-٨) بيانات الدول الأربع. ويبدو من هذا الشكل أنه من الصعب تعيين منوال واضح من بيانات كل من بلجيكا ولكسمبرج بصفة خاصة، لوجود أكثر من فئة منوالية واحدة.

مزايا ومثالب مقاييس النزعة المركزية:

بعد ذلك العرض التفصيلى لطرق حساب كل مقياس من مقاييس النزعة المركزية الثلاثة وبعد معرفة العلاقة بينها، فإنه يمكننا الآن معرفة دلالة الفروق بين هذه المقاييس فى الوصف الإحصائى واستنباط مزاياها وعيوبها كأدوات للقياس فى التحليل الكمى للبيانات الجغرافية. فالمتوسط الحسابى كأحد مقاييس النزعة المركزية يكتسب أهمية خاصة بين الأنواع الثلاثة للمتوسطات باعتباره مقياساً دقيقاً يقوم على أسس رياضية سليمة، كما أن له مميزات وخصاءص تسمح باستخدامه فى إيجاد مقاييس أخرى نحتاج إليها فى التحليل الاحصائى للبيانات. ومن بين هذه الخصائص أن مجموع انحرافات قيم المفردات عن المتوسط الحسابى، دون غيره من المقياسين الآخرين (الوسيط والمنوال)، يساوى صفراً، ويمكن اثبات هذه الخاصية جبرياً فى الخطوات التالية:

الإنحراف عن المتوسط = (س - سَ) ويرمز له بالرمز ح

$\therefore$ ح$_1$ = (س$_1$ - سَ)

ح$_2$ = (س$_2$ - سَ)

•

•

$ح_ن = (س_ن - \bar{س})$

وبالجمع نجد أن:

$مجـ ح = مجـ \left[(س - \bar{س}) + (س - \bar{س}) + ... + (س - \bar{س}) \right]$

$= مجـ س - ن \bar{س}$

وبالتعويض عن قيمة $\bar{س}$ (المتوسط الحسابى) بالصيغة $\frac{مجـ س}{س}$

$\therefore$ $مجـ ح = مجـ س - ن$

$= مجـ س - مجـ س = صفر$

وللمتوسط الحسابى مميزات أخرى منها أن مجموع مربع الإنحرافات عن المتوسط يقل عن مجموع مربع الإنحرافات عن أى قيمة أخرى، وهذه خاصية هامة ستستخدم فى حساب مقاييس التشتت فيما بعد. كما أنه إذا كان لدينا عدداً من أزواج القيم لمتغيرين مستقلين فإن المتوسط الحسابى لمجموع قيم المتغيرين يساوى مجموع المتوسطين الحسابيين لهذين المتغيرين. ويمكن إثبات ذلك على النحو التالى:

$ت_١ = أ_١ + ب_٢$

$ت_٢ = أ_١ + ب_٢$

. .
. .
. .
. .

وبالجمع نجد أن

مجـ ت = مجـ أ + مجـ ب

وبالقسمة على عدد المفردات ن نجد أن:

$$\bar{ت} = \bar{أ} + \bar{ب}$$

حيث تَ هي المتوسط الحسابي لمجموع المتغيرين، أَ هي المتوسط الحسابي للمتغير الأول، بَ هي المتوسط الحسابي للمتغير الثاني.

وكذلك عند إضافة قيمة ثابتة «ك» إلى كل قيمة من قيم مفردات البيانات أو إذا ضربت كل قيم المفردات في قيمة ثابتة «ك»، فإن المتوسط الحسابي في الحالة الأولى يساوي المتوسط الحسابي للمفردات قبل الإضافة مضافاً إليه القيمة الثابتة بينما في الحالة الثانية يساوي المتوسط الحسابي للبيانات قبل الضرب مضروباً في القيمة الثابتة. ويمكن أن نعبر عن ذلك بالرمز كمايلي:

$$\text{متوسط (س + ك)} = \frac{\text{مجـ س + ن ك}}{\text{ن}}$$

$$= \bar{\text{س}} + \text{ك}$$

$$\text{ومتوسط (س} \times \text{ك)} = \frac{\text{مجـ س + ك}}{\text{ن}}$$

$$= \bar{\text{س}} \times \text{ك}$$

أما عيوب المتوسط الحسابي فتنحصر في أنه إذا إحتوت مجموعة من قيم المفردات على بعض القيم الشاذة أو المتطرفة (القيم الكبيرة جداً أو الصغيرة جداً) فإن المتوسط الحسابي للمجموعة يكون في هذه الحالة مضللاً، وذلك لأن حساب المتوسط سيتأثر بهذه القيم على الرغم من قلة عددها بين مجموعة القيم، وعندئذ يفضل استخدام مقياس آخر من مقاييس النزعة المركزية لوصف هذه المجموعة. وفي التوزيعات التكرارية المتماثلة (شكل ٤-٥أ)، ولا يبدو هذا العيب في المتوسط الحسابي جوهرياً لعدم وجود اختلافات كبيرة بين قيم المجموعة، فنصفها يكون أكبر من المتوسط بينما النصف الآخر دونه. أما في التوزيعات الملتوية فإن الوضع يختلف عند حساب المتوسط الحسابي، إذ تؤثر القيم المتطوفة في مثل هذه التوزيعات تأثير كبيراً على قيمته. ويمكن بيان ذلك المثال التالي الذى يوضح كمية

الأمطار السنوية التى سجلها أحد المراصد فى إقليم جاف خلال عشرة أعوام متتالية (شكل رقم ٤-٩).

كمية المطر السنوى (بالبوصة) : صفر، ١، صفر، صفر، ١٠، ٢، ٢٥، صفر، صفر، ٢.

فإذا حسبنا المتوسط الحسابى للبيانات السابقة نجد أن.

$$\bar{س} = \frac{مجـ س}{ن} = \frac{٤٠}{١٠} = ٤ \text{ بوصات}$$

وبالتالى يمكن ملاحظة أن كمية الأمطار السنوية قد زادت عن متوسطها الحسابى (٤ بوصات) مرتين فقط (عندما كانت ١٠، ٢٥ بوصة) فى السنتين المطيرتين خلال الأعوام العشرة، بينما قلت كمية الأمطار عن هذا المتوسط فى الثمانى سنوات الباقية. ومن هذا نرى كيف أن العاملين المطيرين (٣٥ بوصة أو ٨٧,٥٪ من كمية أمطار السنوات العشر) قد ساهما فى رفع قيمة المتوسط الحسابى وأثرا فيها أكثر مما أثرت به السنوات الأربع الأخرى التى لم تسقط بها أمطار على الإطلاق. وفى هذا المثال فإن استخدام المتوسط الحسابى كمقياس للموضع يكون مضللا حيث أن التوزيع التكرارى لهذه البيانات توزيعاً ملتوياً كما هو واضح فى الشكل رقم (٤-٩)، ويفضل فى مثل هذه الحالة استخدام مقياس آخر من مقاييس الموضع. وفى هذا المثال المحدد فإن الوسيط يساوى ٠,٥ بوصة والمنوال يساوى صفر بوصة، وكلاهما بلا شك يعطى مؤشراً أو ملخصاً للصورة العامة للنظر فى المناطق الجافة تفوق ما يعطيه المتوسط الحسابى. وبذلك يجب الحذر الشديد عند حساب المتوسط الحسابى لأى مجموعة من البيانات التى فيها قيم شديد التطرف فى الكبر والصغر إذ يعجز متوسطها الحسابى عن إعطاء الصورة المطلوبة.

ومن عيوب المتوسط الحسابى، بالإضافة إلى ما سبق، صعوبة حسابه من جداول التوزيع التكرارى غير المنتظم أو من جداول التوزيعات المفتوحة، وهى

التوزيعات التكرارية غير المعروف الحد الأدنى للفئة الأولى بها أو الحد الأعلى للفئة الأخيرة أو كليهما معاً. ويفضل أيضاً فى الحالة الأخيرة استخدام مقياس آخر من مقاييس النزعة المركزية يكون أكثر ملائمة لمثل هذه الجداول المفتوحة. وأخيراً فإن المتوسط الحسابى يصعب إيجاده بالطرق البيانية متماثلاً فإن المتوسط الحسابى يقع على المحور الأفقى عند تلاقيه مع محور التماثل الرأسى بينما فى التوزيعات البسيطة الالتواء يقع المتوسط الحسابى قريباً من أكبر التكرارات من ناحية الذيل الأطوال للتوزيع والذى يلاحظ منه تأثر المتوسط بالقيم المتطرفة فتجعله يتحيز لها (شكل ٤-٥أ، ب، جـ).

وعلى الرغم من كل عيوب ومثالب المتوسط الحسابي السابق ذكرها فإنه كمؤشر ومقياس إحصائى له أهمية كبيرة كما يعد مفيداً بل ضرورة فى عملية البحث والتحليل الكمى فى الدراسة الجغرافية.

أما الوسيط فقد سبق أن عرفنا أنه يستخدم كقيمة متوسطة، لأنه يمثل أحياناً المتوسط المتوقع Mean Expectation أو ينوب عنه، حيث يكون عدد المفردات التى تسبقه مساوية تماماً لعدد المفردات التى تليه. كما تكون قيم المفردات لها نفس الأهمية عند حساب الوسيط بصرف النظر عن حجمها كبيرة أو متوسطة أو صغيرة) فيما عدا القيم القيمة الوسيطية (المركزية) إذا كان عدد المفردات فردياً، أو القيمتين المتتابعتين المركزتين فى حالة إذا كانت المفردات عدداً زوجياً. ومعنى ذلك أننا قد نجد عدة مجموعات متباينة من البيانات التى تختلف بدرجة كبيرة فى مجموع قيمها ولكنها تشترك فى قيمة الوسيط. وهذا يدل على أن الوسيط الحسابية الأخرى إلا بأسلوب عام. وهو بهذا يعانى من نفس مثالب المنوال التى سيرد ذكرها بعد قليل، ومع ذلك فإن للوسيط خاصية هامة ترتبط بوضوح تعريفه ذلك لأن موقعه النسبى بين المفردات أو بين تكرار حدوثها يساعد على تحديده بسهولة. هذا بالإضافة إلى أن الوسيط يتميز بأنه أقل تأثرا بالقيم المتطرفة عن المتوسط الحسابى حيث لا تدخل هذه القيم فى حسابه، كما أنه يمكن حسابه من التوزيعات المفتوحة حين يتعذر حساب المتوسط الحسابى لها، وهو بذلك يعد

مقياساً إحصائياً مفيداً كأسلوب توضيحي يفوق المتوسط الحسابي والمنوال، وما لذلك من أهمية كبيرة في كثير من مجالات التحليل الإحصائي للبيانات الجغرافية.

على أنه في بعض الحالات يواجه حساب الوسيط بعض الصعوبات. كما في حالة المتغيرات الوثابة (غير المتصلة) ذات القيم المزدوجة. فمثلاً إذا كان لدينا عدد سكان ستة منازل بأحد الشوارع هو: ١٢١ . ١٣٤ ، ١٥١ ، ١٦٢ ، ١٧٧ ، ١١٨ فإن الوسيط حسب التعريف السابق يكون:

$$\text{الوسيط} = \frac{١٦٢ + ١٥١}{٢} = \frac{٣١٣}{٢} = ١٥٦,٥$$

وهذه القيمة لا وجود لها إذ أن صفة المتغير أنه وثاب أي لا يأخذ قيما كسرية، فلا معنى إذن لعدد سكان ١٥٦,٥ شخصاً. كما توجد صعوبة أيضاً في تحديد الوسيط في المتغيرات المتصلة ذات القيم القليلة العدد أو التي بينها قيم تتكرر بكثرة عن غيرها. فإذا كان لدينا أطوال خمسة روافد نهرية (بالمتر) هي: ١٦٢ ، ١٦٢ ، ١٦٢ ، ١٦٥ ، ١٦٧ فإن القيمة الوسيط هي ١٦٢ متراً ولا توجد أي قيمة أصغر منها وأن ٤٠ ٪ من القيم فقط أكبر منها. وبناء على ذلك فإنه كلما كان عدد مفردات البيانات صغيراً كلما كان من الأفضل استخدام مقياساً آخر غير الوسيط لوصف هذه البيانات.

ومن التعريفات السابق شرحها لمقاييس النزعة المركزية عرفنا أن المنوال بمعناه الدقيق يحدد القيم لأكثر شيوعاً وتكراراً في التوزيع، ولكن قد يصعب أحياناً تحديد موقع (قيمة) المنوال بدقة داخل التوزيع. وتظهر هذه الصعوبة في إحدى الحالتين الآتيين: أولهما، إذا كان التوزيع يحتوى على أكثر من قيمة متساوية في تكرارها أي أكثر من منوال واحد أو قد لا يكون به منوال على الإطلاق. ففي بيانات الجدول رقم (٤-١٢) الذى يوضح انتاج الحديد الخام لبعض دول غرب أوربا لاحظنا صعوبة تحديد منوال واضح في حالة البيانات الخاصة لكل من بلجيكا

ولكسمبرج حيث توجد فئتان من التكرارات المتساوية أو منوالان فى كل مجموعة عريضة من البيانات (شكل رقم ٤-٨). وتبدو الصعوبة الثانية لتحديد قيمة المنوال فى إختيار فئات التوزيع التكرارى لبيانات جدول كمية الأمطار السنوية السابق عرضه (جدول رقم ٤-١١) إلى فئات (٢٢•٢٣,٩٩)، (٢٤-٢٥,٩٩)، (٢٦-٢٧,٩٩) ... الخ بدلاً من (٢١- ٢٢,٩٩)، (٢٣- ٢٤,٩٩). (٢٥-٢٦,٩٩).... الخ فيصبح جدول التوزيع التكرارى كمايلى:

الفئات (بالبوصة)	عدد مرات الحدوث (التكرار)
٢٢ - ٢٣,٩٩	١
٢٤ - ٢٥,٩٩	٩
٢٦ - ٢٧,٩٩	٣
٢٨ - ٢٩,٩٩	٨
٣٠ - ٣١,٩٩	٤
٣٢ - ٣٣,٩٩	٢
٣٤ - ٥٣,٩٩	٢
٣٦ - ٣٧,٩٩	١

وفى مثل هذا التوزيع تصبح الفئة المتوالية (٢٤- ٢٥,٩٩) بدلاً من الفئة (٢٥ - ٢٦٧,٩٩) وستصبح قيمة المنوال ٢٥,١٤ بوصة بدلاً من ٢٦,٣٣ بوصة فى التوزيع السابق، كما يصبح التوزيع فى هذه الحالة مزدوج المنوال Bimodal وتعنى هذه الصعوبة أن اختيار المنوال كمقياس للموضع ليس اختياراً موفقاً، كما أنه عملياً يعد شكلاً غير دقيق للقيمة المتوسطة، وربما تنتج فى جزء منها عن الاختيار الذاتى لفئات التوزيع التكرارى كما رأينا. وعلاوة على ذلك فإن المنوال لا يحتوى على خصائص رياضية حقيقية بل له على أحسن الفروض علاقة عامة

بالمتوسط الحسابى فقط ولكن لدرجة لا تمكن من استخدامه فى صيغة تشتق منها صفات أخرى للبيانات. وفيما عدا بعض الأغراض البيانية واستخدامه فى بعض العمليات الحسابية العامة، فإنه لا يصلح لأن يؤخذ مقياساً للموضع فى الدراسات الجغرافية كما أنه لا يوصى باستخدامه بدرجة كبيرة فى التحليل الاحصائى للبيانات الجغرافية.

من كل ماسبق وحتى نتمكن من معرفة أهمية وفائدة كل مقياس من مقاييس النزعة المركزية الثلاثة، ينبغى أن نعرف نمط التوزيع التكرارى الذى تلخصه هذه المقاييس فى قيمة واحدة، أى معرفة الأحوال الفعلية للبيانات وكيفية انتشارها حول القيم المتوسطة. ومن الأساليب الإحصائية التى يعتمد عليها فى قياس هذه الخاصية مايعرف بمقاييس الانحراف (التشتت) والاختلاف، وهذا ما سنشرحه فى الفصل التالى.



الفصل الخامس

مقاييس الإنحراف (التشتت) والإختلاف

Measures of Deviation (Dispersion) and Variability

مقاييس الإنحراف (التشتت) والإختلاف

تكلمنا فى الفصل السابق عن المقاييس المختلفة للنزعة المركزية (المتوسطات) وكيفية قياس كل منها، وعرفنا أنها تهدف إلى تبسيط وتحديد الاتجاه العام لمجموعة من قيم مفردات البيانات عن طريق تلخيصها فى قيمة متوسطة مركزية واحدة تصف أو تمثل قيم هذه المجموعة وهى ما أطلقنا عليه صفة «التركز» فى البيانات. وكما لاحظنا فإن مقاييس النزعة ركزية لاتكفى وحدها لإعطاء فكرة دقيقة عن مجموعة من البيانات وصفها وصفاً كاملاً إذ أنها لاتبين طبيعة هذه البيانات ولا كيفية توزيع مفرداتها، كما أن استخدامها فقط لمقارنة عدة مجموعات من البيانات لايكفى لإظهار حقيقة المقارنة. وعلى ذلك فالوصف الدقيق لمجموعة من القيم الأصلية للبيانات أو مقارنة هذه المجموعة بمجموعة أو عدة مجموعات أخرى بدقة يجب ألا يقتصر على وصف أو مقارنة متوسطاتها، بل يجب الوقوف على صفات كل مجموعة. ويتحقق ذلك بوصف درجة الإختلاف بين قيم مفردات المجموعة وتباعدها عن قيمتها المتوسطة، أو بعبارة أخرى وصف درجة إنحراف أو تشتت هذه القيم.

والمقصود إحصائياً بالانحراف أو التشتت هو مدى تباعد وتناثر (إنتشار Scat-ter) قيم مفردات البيانات عن بعضها البعض. فإذا كانت قيم المفردات متقاربة من بعضها البعض (أو إذا تساوت جميع القيم) فإن مدى التناثر يكون صغيراً وبالتالى

يدل على تجانس هذه القيم، أما إذا كانت القيم منتشرة فيما بينها أى متباعدة عن بعضها البعض فإن مدى التناثر يكون كبيراً، ويتخذ ذلك دليلاً على عدم التجانس بين قيم المفردات. وعلى ذلك يمكننا إتخاذ مقدار إنحراف أو تشتت القيم كمقياس لتجانس البيانات أو كمقياس لتركز القيم وقربها من بعضها أو لتبعثرها وتباعدها بعضها عن البعض. وتقاس درجة التجانس أو التركز ببعض المقاييس التى تعرف باسم مقاييس الإنحراف (التشتت) والاختلاف، بينما تقاس إتجاهات التركز والتطرف فى القيم بما يسمى بمقاييس الإلتواء Skewness ومقاييس التفرطح Kurtosis.

أنواع الإنحراف:

إذا أردنا توضيح شكل التشتت لمفردات المتغيرات موضع الدراسة بيانياً، فإن أبسط وأنسب وسيلة لتحقيق ذلك هو ما يتم بواسطة منحنى التوزيع التكرارى Frequency Distribution Curve الذى سبق شرح وتطبيق طريقة تمثيله، بينما «شكل الانتشار Scatter Diagram» الذى تتحرك فيه قيم المتغيرات لايفيد كثيراً فى بيان شكل التشتت بقدر ما يفيد فى توضيح نوع ودرجة العلاقة الارتباطية بين المتغيرات (سبق شرح هذا النوع من الرسوم عند دراسة طرق العرض البيانى للبيانات الاحصائية).

أما إذا أردنا التعبير عن التشتت لمجموعة من قيم المفردات بصورة رقمية فإن ذلك يتم عن طريق مقارنة أقل القيم بأكبرها فى قيم المجموعة الواحدة أو بما يعرف بالمدى المطلق Range. فعلى سبيل المثال، إذا كانت لدينا عينتين لأطوال خمسة من الأودية النهرية (بالكيلو متر) فى منطقتين مختلفين تتوزع قيمها على النحو التالى:

العينة الأولى: ١٧، ١٤، ١٠، ٧، ٢ (المتوسط = ١٠، الوسيط = ١٠)

العينة الثانية: ١٢، ١١، ١٠، ٩، ٨ (المتوسط = ١٠، الوسيط = ١٠)

فنلاحظ أن قيمة المتوسط الحسابى والوسيط لكل من هاتين العينتين متساوية أى أن العينتين تشتركان فى أكثر من مقياس من مقاييس الموضع ومع ذلك نجد أن هناك فرقاً كبيراً بين العينتين من حيث طبيعة توزيع قيم المفردات فى كل منهما، إذ تتوزع أطوال أودية العينة الأولى فى مدى تبعثر أكبر (من ٢ - ١٧ كيلو متراً) بينما تتوزع أطوال أودية العينة الثانية فى مدى تبعثر أقل (من ٨ - ١٢ كيلو متراً). ومعنى ذلك أن التشتت بين مفردات العينة الأولى أكبر منه بين مفردات العينة الثانية (أى أن قيم العينة الأولى أقل تجانساً من العينة الثانية). وكذلك نلاحظ أن هناك إختلاف واضح فى انحراف قيم المفردات عن المتوسط الحسابى فى كل عينة إذ ينحصر الانحراف عن المتوسط للعينة الأولى بين +٧ و -٨ بينما ينحصر فى العينة الثانية بين +٢ و -٢.

ومهما يكن من أمر فإن المدى كمقياس للتشتت، رغم بساطته، لايعد دقيقاً للتعبير من وصف مجموعة من قيم المفردات بالتجانس (التشتت) فهو لايعطى فكرة جيدة عن مدى تباعد القيم عن بعضها البعض، وربما يكون مضللاً فى الحالات التى يوجد فيها قيمة متطرفة (شاذة) فى البيانات مما يسبب زيادة كبيرة فى المدى الذى يستدل منه على أن قيم المفردات غير متجانسة ومشتتة تشتتاً كبيراً بينما تكون القيم كلها - ما عدا القيمة المتطرفة - متقاربة. كما أن المدى لايعطى ملخصاً كافياً للتشتت أو التبعثر فى قيم المفردات، أو لايوجز التشتت فى قيمة واحدة يعتمد عليها فى التحليل الاحصائى للبيانات. وبناء على ذلك هناك ثلاثة مقاييس أخرى أكثر دقة وفائدة من حيث تلخيص التشتت وإنتشار ومفردات البيانات ذلك لأنها تهدف إلى، أو ينتج عنها، قيم رقمية تعرف باسم «القيم الانحرافية Deviation Values»، وهى عبارة عن إختلافات أو فروق قيم مفردات المجموعة عن المتوسط. وهذه المقاييس هى: الانحراف الربيعى، الانحراف المتوسط، للتباين (والانحراف المعيارى). والمقياس الأول من هذه المقاييس يستخدم الوسيط، بينما يستخدم المتوسط الحسابى فى المقياسين الآخرين. أما المنوال فلا يستخدم فى

حالة قياس التشتت كبديل للمتوسط لأنه لايظهر القيم المنحرفة بشكل فعال، وهذه ولاشك واحدة أخرى من مساوئ المنوال التى تقلل بدرجة كبيرة جداً إستخدامه فى التحليل الجغرافى الكمى. وفيما يلى شرح للمقاييس الثلاثة كل على حدة وبيان حساب التشتت بواسطتها.

١ - الانحراف الربيعى Quartile Deviation

يستخدم الانحراف الربيعى أو ما يسمى بنصف المدى الربيعى - Semi Inter quartile Range لقياس التشتت فى البيانات التى تتميز بتطرف بعض قيم مفرداتها، ويهدف إلى التغلب على أهم عيوب المدى المطلق - وهو إعتماده على القيم المتطرفة، وذلك بحذف الربع الأول والربع الأخير من قيم المتغير موضع الدراسة وحساب المدى للقيم الأخرى. أو بعبارة أخرى يحدد النصف الأوسط لمجموعة قيم المتغير بعد ترتيبها حسب قيمتها ويستخرج المدى الذى تقع فيه القيم الوسيطة، وكلما كان المدى الذى ينتشر عليه هذا النصف من القيم كبيراً كان التشتت كبيراً بين قيم المتغير والعكس يكون صحيحاً، وبذلك يمكن الحصول على مقياس للتشتت يعتبر أفضل من مقياس المدى المطلق.

وتجدر الإشارة إلى أن طريقة إيجاد الانحراف الربيعى لمجموعة من القيم تشبه الطريقة التى يوجد بها الوسيط، ولذلك فإن مزايا ومثالب كل منهما واحدة. فالوسيط كما سبق أن عرفنا هو القيمة التى تقسم قيم المجموعة المرتبة حسب قيمتها (تصاعدياً أو تنازلياً) إلى نصفين متساويين فى العدد. ويمكن قياس التشتت عن القيمة الوسيطية بقسمة كل نصف المجموعة إلى نصفين آخرين، وبذلك تصبح المجموعة مقسمة إلى أربعة أقسام متساوية - ويسمى كل حد من حدود التقسيم «بالربيع»، ويكون عدد الحدود الفاصلة فى هذه الحالة ثلاثة: يطلق على الحدين الفاصلين الأول والأخير بالربيع الأدنى Lower Quartile والربيع الأعلى Upper Quartile على الترتيب، بينما الحد الثالث (الأوسط) هو الوسيط. ويظهر ذلك فى الشكل التالى (شكل رقم: ٥-١) الذى يمثل بيانات كمية المطر السنوى التى

١	٣٦,٥٠	النهاية العظمى
٢	٣٤,٨١	
٣	٣٤,٤٢	
٤	٣٣,٣٤	
٥	٣٢,٨٧	
٦	٣١,٩٣	
٧	٣٠,٩٢	
٨	٣٠,٩٥	الربيع الأعلى
٩	٣٠,١٧	
١٠	٢٩,١٢	
١١	٢٩,١١	
١٢	٢٨,٩٥	
١٣	٢٨,٩٠	
١٤	٢٨,٥٩	
١٥	٢٨,٤٥	
		الوسيط (٢٨,٢٧)
١٦	٢٨,٠٨	
١٧	٢٨,٠٠	
١٨	٢٦,٨٣	
١٩	٢٦,٥٧	
٢٠	٢٦,٠٢	
٢١	٢٥,٩٧	
٢٢	٢٥,٣٨	
٢٣	٢٥,٥٧	الربيع الأدنى
٢٤	٢٥,٢٧	
٢٥	٢٥,١٩	
٢٦	٢٥,١٨	
٢٧	٢٥,١٠	
٢٨	٢٤,٨٧	
٢٩	٢٤,٠١	
٣٠	٢٢,٤٧	النهاية الصغرى

شكل رقم (٥-١)

الوسيط والربيعين (الأدنى والأعلى) لكميات المطر السنوية المسجلة بمرصد Bidston - انجلترا فى الفترة من ١٩٠١ - ١٩٣٠

سجلها مرصد Bidston – إنجلترا فى الفترة من ١٩٠١ – ١٩٣٠ (جدول رقم ٤ – ١٠).

من الشكل رقم (٥–١) يتضح أن الربيع الأدنى يفصل ٢٥٪ من المفردات التى تشمل أقل القيم عن بقية المفردات، بينما يفصل الربيع الأعلى ٢٥٪ من المفردات التى تحتوى على أعلى القيم. أى أن حدودهما تحصر القيم الوسطى (أو ٥٠٪ من المفردات) التى تتشتت حول قيمة الوسيط. ومن الشكل أيضاً نرى أن قيمة الوسيط ٢٨,٢٧ بوصة، وهى نفس القيمة التى حصلنا عليها سابقاً عند تطبيق طريقة حساب الوسيط لهذا المثال، أما قيمة الربيع الأدنى فهى عبارة عن القيمة الوسيطية لنصف عدد القيم أسفل الوسيط، أى أنها القيمة الثامنة (من أسفل) من بين الخمس عشرة قيمة التى تقع أسفل الوسيط والتى تساوى ٢٥,٥٧ بوصة. وبنفس الطريقة نحصل على قيمة الربيع الأعلى وهى القيمة الثامنة (من أعلى) من بين الخمس عشرة قيمة التى تقع فوق الوسيط والتى تساوى ٣٠,٥٩ بوصة. ويعرف الفرق بين قيمتى الربيع الأعلى والربيع الأدنى وهو ٣٠,٥٩ – ٢٥,٥٧ = ٥,٠٢ بوصة بالمدى الربيعى. وهذا المدى يقع منحرفاً عن الوسيط، ولكن فى الحالات التى يكون فيها منحنى التوزيع التكرارى للبيانات متماثلاً ومتوازناً تماماً (منحنى معتدل) فإن كلاً من الربيعين سيقع فى منتصف المسافة بين الوسيط والقيم النهائية (العليا والسفلى) للتوزيع. أى أن كلاً منهما سيبعد عن الوسيط بنصف قيمة المدى وهى ٥,٠٢ ÷ ٢ = ٢,٥١ بوصة للمثال بين أيدينا. وهذه القيمة تعرف بالإنحراف الربيعى أو نصف المدى الربيع وهى التى تتخذ كدليل أو مؤشر يبين مقدار مدى قيم النصف الأوسط حول الوسيط لمجموعة من المفردات المرتبة حسب قيمتها. ويمكن أن نعبر عن ذلك بالصيغة الآتية:

$$\text{الانحراف الربيعى} = \frac{\text{قيمة الربيع الأعلى} - \text{قيمة الربيع الأدنى}}{٢} \quad ...(٥ - ١)$$

ايجاد الانحراف الربيعى للبيانات المبوبة:

لحساب الانحراف الربيعى من بيانات مبوبة أى من جداول التوزيعات التكرارية تتبع الخطوات التالية:

١- تحسب من جدول التوزيع التكرارى التوزيع التكرارى المتجمع الصاعد للتكرارات.

٢- توجد ترتيب كل من الربيعين الأدنى (ن/٤) والأعلى (ن × ٣/٤).

٣- تعين قيمة كل من الربيعين باستخدام الصيغة الآتية:

قيمة الربيع = الحد الأدنى للفئة الربيعية + ^

(ترتيب الربيع − التكرار المتجمع السابق للفئة الربيعية / التكرار الأصلى للفئة الربيعية) × طول الفئة

ويمكن بيان ذلك المثال السابق الخاص بكميات المطر السنوية المسجلة بمرصد Bidston إنجلترا فى الفترة من ١٩٠١ – ١٩٣٠.

جدول رقم (٥-١)

كمية الأمطار السنوية المسجلة بمرصد Bidston

انجلترا فى الفترة من ١٩٠١ - ١٩٣٠

الفئـــــات (بوصة)	التكـــــرار	التكرار المتجمع الصاعد	
		الحدود العليا للفئات	التكرار المتجمع الصاعد
٢١ - ٢٢,٩٩	١	أقل من ٢٣	١
٢٣ - ٢٤,٩٩	٢	أقل من ٢٥	٣
٢٥ - ٢٦,٩٩	١٠	أقل من ٢٧	١٣
٢٧ - ٢٨,٩٩	٦	أقل من ٢٩	١٩
٢٩ - ٣٠,٩٩	٥	أقل من ٣١	٢٤
٣١ - ٣٢,٩٩	٢	أقل من ٣٣	٢٦
٣٣ - ٣٤,٩٩	٣	أقل من ٣٥	٢٩
٣٥ - ٣٦,٩٩	١	أقل من ٣٧	٣٠
المجموع	٣٠		

ويكون ترتيب الربيع الأدنى = $\frac{٣٠}{٤}$ = ٧,٥

∴ قيمـــة الربيــع الأدنـى = ٢٥٠٠ + $\left(\frac{٧,٥ - ٣}{١٠}\right)$ × ٢

= ٢٥,٠ + $\frac{٤,٥ \times ٢}{١٠}$

= ٢٥,٠ + ,٩ = ٢٥,٩ بوصة

وترتيب الربيع الأعلى = $\frac{٣٠ \times ٣}{٤}$ = ٢٢,٥

∴ قيمة الربيع الأعلى = ٢٩,٠ + $\left(\frac{٢٢,٥ - ١٩}{٥}\right) \times ٢$

= ٢٩,٠ + $\frac{٣,٥ \times ٢}{٥}$

= ٢٩,٠ + ١,٤ = ٣٠,٤ بوصة

ويكون الانحراف الربيعى أو (نصف المدى الربيعى) = $\frac{\text{قيمة الربيع الأعلى} - \text{قيمة الربيع الأدنى}}{٢}$

= $\frac{٣٠,٤ - ٢٥,٩}{٢}$

= $\frac{٤,٥}{٢}$ = ٢,٢٥ بوصة

ومن الواضح أن نصف عدد التكرارات (سنوات المطر) ينحصر بين قيمتى يعين الأدنى والأعلى (أى الكميتين ٢٥,٩ و ٣٠,٤ بوصة)، وهذا النصف سط الذى لا يتأثر بالقيم المتطرفة، كما أنه يمثل ٥٠٪ من البيانات التى تتشتت ها عن الوسيط بمقدار ± نصف المدى الربيعى. وبناء على ذلك يمكن وصف حراف الربيعى بأنه معدل الانحراف المتوقع عن القيمة الوسيطية أو بمعنى آخر نصف القيم تختلف عن الوسيط بأكثر من مقدار نصف المدى الربيعى، بينما ف الآخر يختلف عنه بأقل من هذا المقدار.

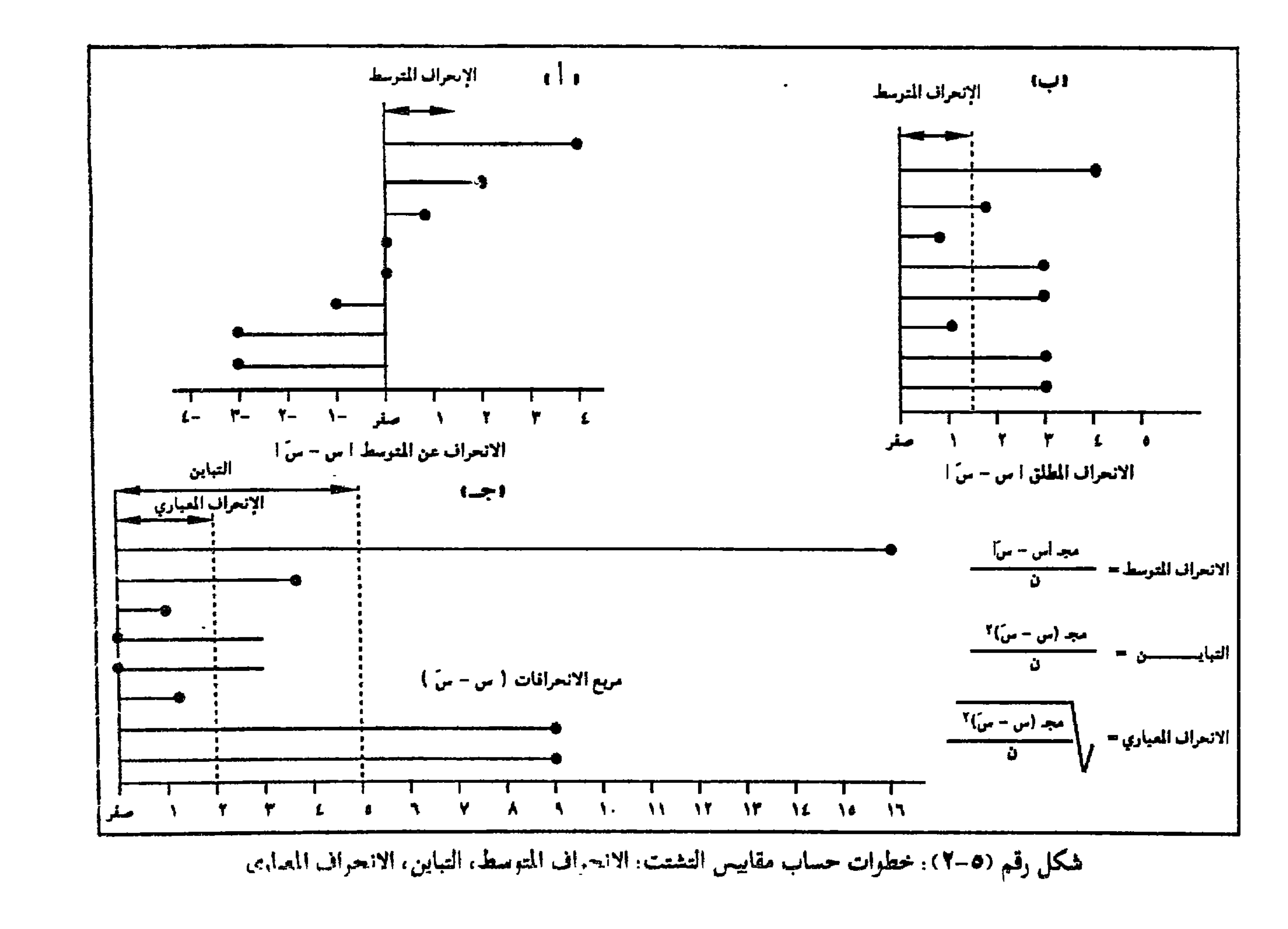
ومن الملاحظ أيضاً أنه إذا قارنا الانحراف الربيعى (٢,٢٥) وقيمة المدى المطلق ٣ - ٢١ = ١٦ بوصة) وهما يقيسان التشتت لنفس سنوات المطر - نجد أن تهما مختلفتان، وهذا أمر متوقع فكل منهما يقوم على أساس مختلف. ولذلك يتحتم عند إجراء مقارنة بين تشتت عدة مجموعات من البيانات أن يكون ى التشتت بطريقة واحدة لكل المجموعات حتى تكون المقارنة صحيحة. وعلى

أية حال، فإن الانحراف الربيعى أو نصف المدى الربيعى لايعد مقياساً دقيقاً لتشتت البيانات بل يعتبر مقياساً لتوضيح مدى تشتت البيانات بعداً أو قرباً من مقياس الموضع (الوسيط). ومع ذلك فهو وسيلة هامة لبيان مدى تشتت البيانات حول نقطة معينة، ولذلك فهو يستخدم كثيراً فى تحليل البيانات المناخية، أو لإبراز المعلومات الاقتصادية وتحليلها - فى مجال الجغرافية الاقتصادية - وبصفة خاصة عند دراسة وتحليل التوطن الصناعى فى أحد الأقاليم الجغرافية.

الانحراف المتوسط: Mean Deviation

لاحظنا من حساب كل من المدى المطلق والانحراف الربيعى أن كلاً منهما لايعطى صورة كاملة عن التشتت فى قيم البيانات. فالمدى يعتمد على القيم الطرفية المحددة للتوزيع، التى غالباً ما تكون متطرفة، ولايهتم بتوزيع بقية القيم، بينما يهمل الانحراف الربيعى نصف عدد للقيم فى أطراف توزيع البيانات. وكما سبق القول أن القياس الصحيح للتشتت يقوم على أساس مدى تباعد القيم عن بعضها فإذا كانت القيم قريبة من بعضها فإنه تكون متجمعة (متركزة) حول قيمة فى الوسط وكلما كانت متناثرة (مبعثرة) كلما تباعدت عن هذه القيمة. وبناء على ذلك فإننا نلجأ إلى قياس التشتت على أساس بين القيم المختلفة للبيانات وقيمة متوسطة مثل المتوسط الحسابى (الذى يعتبر أكثر مقاييس النزعة المركزية حساسية اللتغيرات فى أى قيمة من القيم).

ولما كان من المعروف أن مجموع انحرافات القيم عن متوسطها الحسابى يساوى صفراً وذلك لأن مجموع الانحرافات الموجبة عن المتوسط الحسابى يساوى مجموع الانحرافات السالبة عنه، ولذا فإنه لايمكن أخذ مجموع الانحرافات عن المتوسط الحسابى كمقياس للتشتت، ولكن إذا أهملت إشارات الانحرافات (انحراف موجب أو انحراف سالبة) وجمعنا الانحرافات المجردة من الإشارة ثم قسمنا هذا المجموع على عدد القيم فإننا نحصل بذلك على مقياس للتشتت يطلق عليه اسم الانحراف المتوسط (شكل رقم ٥-٢). ويمكن تلخيص ذلك فى الصيغة الجبرية الآتية:



شكل رقم (٥-٢) : خطوات حساب مقاييس التشتت: الانحراف المتوسط، التباين، الانحراف المعياري

$$\text{الانحراف المتوسط} = \frac{\text{مجـ} |ح|}{ن} \quad \ldots \quad \ldots \quad \ldots \quad (٥ - ٢)$$

حيث مجـ هى المجموع، ن هى عدد القيم، ح هى الانحراف عن المتوسط الحسابى، ويرمز الخطان الرأسيان إلى أن الانحرافات تؤخذ بدون إشاراتها الجبرية (أى الانحرافات المطلقة). وتمتاز قيمة الانحراف المتوسط بأنها تكبر كلما ازداد التشتت أو التفاوت بين القيم وبعدت عن متوسط المجموعة، بينما تقل كلما قل تشتت وانتشار القيم. كما يمتاز الانحراف المتوسط بأنه يأخذ فى الاعتبار جميع قيم المفردات المتاحة.

ويتم قياس التشتت لمجموعة صغيرة من البيانات بواسطة الانحراف المتوسط بأن نوجد أولاً المتوسط الحسابى للقيم (بإحدى الطرق السابق شرحها) ثم نحسب الانحراف (الفرق) بين هذا المتوسط وبين كل قيمة من قيم البيانات ونجمع هذه الانحرافات - مع استبعاد الاشارات - ثم نقسمها على عدد القيم، والمثال التالى يوضح ذلك.

مثال. فى مجموعة البيانات الخاصة بأطوال خمسة من الأودية النهرية فى منطقة ما (بالكليو متر) ١٧، ١٤، ١٠، ٧، ٢ يكون المتوسط الحسابى = $\frac{٥٠}{٥}$ = ١٠ كيلو متراً، وتكون انحرافات القيم عن هذا المتوسط هى + ٧، + ٤، صفر، -٣، -٨ وبإهمال الإشارة السالبة يكون الانحراف المتوسط هو $\frac{+٧ +٤ +\text{صفر} +٣ +٨}{٥}$ = $\frac{٢٢}{٥}$ = ٤,٤ كيلو متراً.

ولإيجاد الانحراف المتوسط لتوزيع تكرارى نحسب المتوسط الحسابى بالطرق العادية ثم نحسب انحرافات مراكز الفئات عن المتوسط الحسابى مع إهمال الإشارة (|ح|)، ثم نوجد حاصل ضرب تكرار كل فئة فى انحراف مركز الفئة مع إهمال الإشارة (|ح| × ك)، ونجمع هذه الانحرافات ونقسم الناتج على عدد التكرارات، وبذلك يكون:

الانحراف المتوسط = $\frac{\text{مجـ |ح| ك}}{\text{مجـ ك}}$ ... ... (٥ -٣)

والجدول الآتى يبين خطوات حساب الانحراف المتوسط لتوزيع القرى فى أحد المراكز حسب الحجم السكانى وذلك على أساس أن المتوسط الحسابى للحجم السكانى للقرى هو ١٩٧٦٤ نسمة.

جدول رقم (٥ - ٢)
الانحراف المتوسط لتوزيع القرى فى أحد المراكز حسب الحجم السكانى لهذه القرى

فئات الحجم (نسمة)	التكرار ك	مراكز الفئات (م)	الانحراف عن المتوسط |ح|	|ح| × ك
١ — ٥٠٠٠	١٢	٢٥٠٠,٥	١٧٢٦٣,٥	٢٠٧٢٦٢,٠
٥٠٠١ — ١٠٠٠٠	١٧	٧٥٠٠,٥	١٢٢٦٣,٥	٢٠٨٤٧٩,٥
١٠٠٠١ — ١٥٠٠٠	١٩	١٢٥٠٠,٥	٧٢٦٣,٥	١٤٨٠٠٦,٥
١٥٠٠١ — ٢٠٠٠٠	١٤	١٧٥٠٠,٥	٩٢٦٣,٥	٣١٦٨٩,٠
٢٠٠٠١ — ٢٥٠٠٠	٧	٢٥٥٠٠,٥	٢٧٣,٥	١٩١٤٨,٥
٢٥٠٠١ — ٣٠٠٠٠	٥	٢٧٥٠٠,٥	٧٧٣,٥	٣٩٦٧٧,٥
٣٠٠٠١ — ٣٥٠٠٠	٥	٣٢٥٠٠,٥	١٢٧٣,٥	٦٣٦٧٧,٥
٣٥٠٠١ — ٤٠٠٠٠	٢	٣٧٥٠٠,٥	١٧٧٣,٥	٣٥٤٧١,٠
٤٠٠٠١ — ٤٥٠٠٠	٥	٤٢٥٠٠,٥	٢٢٧٣٥,٥	١١٣٦٧٧,٥
٤٥٠٠١ — ٥٠٠٠٠	٢	٤٧٥٠٠,٥	٢٧٧٣٥,٥	٥٥٤٧١,٠
٥٠٠٠١ — ٥٥٠٠٠	٣	٥٢٥٠٠,٥	٣٢٧٣٥,٥	٩٨٢٠٦,٥
٥٥٠٠١ — ٦٠٠٠٠	٢	٥٧٥٠٠,٥	٣٧٧٣٥,٥	٧٥٤٧١,٠
٦٠٠٠١ — ٦٥٠٠٠	٢	٦٢٥٠٠,٥	٤٢٧٣٥,٥	٨٥٤٧١,٠
المجموع	٩٥			١١٨١٧٠٨,٥

فيكون الانحراف المتوسط = $\frac{\text{١١٧١٧٠٨,٥}}{\text{٩٥}}$ = ١٢٤٣٩ نسمة تقريباً.

وتجدر الإشارة إلى أن الانحراف المتوسط لايستخدم إلا قليلاً كمقياس للتشتت عند التحليل الكمى للبيانات الجغرافية، ولكننا درسناه هنا كمقدمة لمقياسين آخرين من أكثر مقاييس التشتت إنتشاراً ودقة وهما التباين والانحراف المعيارى.

٣- التباين والانحراف المعيارى: Variance and Standard Deviation

يعد التباين والانحراف المعيارى من أكثر مقاييس التشتت استخداماً فى مختلف البحوث والدراسات فى العلوم الأرضية والاجتماعية، لما لهما من أهمية كبيرة فى الدراسات التحليلية ولاعتماد كثير من المقاييس الاحصائية عليهما. وهما مثل الانحراف المتوسط يعتمدا على استخدام انحرافات القيم عن المتوسط الحسابى. ولكن بدلاً من إهمال الإشارات وللتغلب على الانحرافات السالبة، كما فى حالة الانحراف المتوسط، يتم تربيع الانحرافات فتصبح كلها موجبة ثم يقسم مجموع مربع الانحرافات على عدد هذه الانحرافات فنحصل على متوسط مربع الانحرافات (الانحراف التربيعى المتوسط) أو ما يعرف باسم التباين Variance ويرمز له بالرمز ($ع^2$) فى حالة المجتمع والرمز ($عـ^2$) فى حالة العينة، ويمكن وضع ذلك فى الصيغة الجبرية التالية:

$$\text{تباين المجتمع } (ع^2) = \frac{مجـ (س - مَ)^2}{ن} \quad \dots \quad \dots \quad \dots \quad (٥ - ٤)$$

حيث س هى القيم المعطاة، مَ تشير إلى المتوسط الحسابى للمجتمع و ن هى عدد القيم. ويكون تباين العينة هو.

$$\text{تباين المجتمع } (عـ^2) = \frac{مجـ (س - سَ)^2}{ن} \quad \dots \quad \dots \quad \dots \quad (٥ - ٥)$$

حيث سَ تشير إلى المتوسط الحسابى للعينة. وحيث أن قيمة المتوسط الحسابى للمجتمع (مَ) نادراً ما تكون معروفة فإن تباين العينة يتخذ كتقدير غير متحيز لتباين المجتمع - ويتم ذلك بتصحيح تباين العينة أى بضربه فى عامل التصحيح أو تصحيح «بسل» Besel's Cerrection $\frac{ن}{ن - ١}$ ، ويعرف الناتج «بأحسن تقدير a best estimate لتباين المجتمع ويرمز له بالرمز ($\hat{ع}$)^2 ويمكن التعبير عن ذلك بالصيغة الآتية:

$$(\hat{ع})^{2} = \text{تباين العينة} \times \frac{ن}{ن - ١}$$

$$= \frac{مجـ (س - \bar{س})^{2}}{ن} \times \frac{ن}{ن - ١}$$

$$= \frac{مجـ (س - \bar{س})^{2}}{ن - ١} \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad (٥ - ٦)$$

أى تباين العينة يساوى مجموع مربع الانحرافات عن المتوسط الحسابى للعينة ($\bar{س}$) مقسوماً على عدد الانحرافات المستقلة (ن - ١). وقيمة التباين لاتقل عن الصفر ولكنها تساوى صفراً فى حالة ما إذا كانت جميع قيم مفردات العينة متساوية. وكلما زادت قيمة التباين دل ذلك على زيادة التشتت والاختلاف بين قيم المفردات أو تبعثرها حول المتوسط الحسابى وبالتالى على قلة تجانسها. أما إذا قلت قيمة التباين كان ذلك دليلاً على تركز (تجمع) قيم المفردات حول المتوسط الحسابى وبالتالى زيادة تجانسها. أى أن أمدى تجانس مفردات المجموعة يتناسب عكسياً مع قيمة التباين لها.

مثال:

إذا كانت لدينا البيانات الآتية التى تمثل ثلاث عينات تحتوى كل منها على أعمار ستة أشخاص وأردنا حساب تباين كل مجموعة فستكون خطوات الحساب على النحو المبين بالجدول رقم (٥ - ٣).

وبمقارنة التباين بالانحراف المتوسط - كمقياس للتشتت - للمجموعات السابقة من البيانات والتى تمثل كل مجموعة منها نمطاً مختلفاً للانحرافات عن المتوسط (شكل رقم : ٥ - ٣) نلاحظ أنه على الرغم من أن قيمة الانحراف المتوسط للمجموعات الثلاث واحدة، إلا أنه يمكن القول بأن بيانات المجموعة (حـ) تتسم بأنها أعظم تبايناً وإنتشاراً حول المتوسط من بيانات المجموعة (أ) والمجموعة (ب) بينما تتخذ وضعاً متوسطاً بين المجموعتين الأخريتين. والوسيلة التى يمكن بها تأكيد الانحرافات الكبيرة على حساب الانحرافات الصغيرة هو ما يتم بواسطة تربيع كل الانحرافات ويظهر ذلك جلياً فى بيانات المجموعة (حـ) فى الشكل السابق.

ويعتبر التباين من أشهر مقاييس التشتت (الاختلاف) إلا أنه يقيس مربع الانحرافات. وللحصول على مقياس للتشتت محسوباً بنفس وحدات القيم الأصلية فإننا نستخرج الجذر التربيعى للتباين، وتعرف القيمة الناتجة بالانحراف المعيارى. ويرجع السبب فى أخذ الجذر التربيعى للحصول على هذه القيمة إلى تربيع الانحرافات فى البداية. ولكى تعود الوحدات إلى قيمتها الأصلية بعد التربيع لابد من أخذ الجذر التربيعى ليكون التشتت مقيساً بنفس وحدات القيم الأصلية (بوصات، أمتار، أطنان، سنوات ... إلخ). وبناء على ذلك فإن الانحراف المعيارى للمجتمع يكون:

$$\text{ع} = \sqrt{\frac{\text{مجـ}(\text{س} - \text{م})^{2}}{\text{ن}}} \qquad \ldots \quad \ldots \quad \ldots \quad \ldots (\text{٥} - \text{٧})$$

ويكون الانحراف المعيارى للعينة هو

$$\text{عـ} = \sqrt{\frac{\text{مجـ}(\text{س} - \bar{\text{س}})^{2}}{\text{ن}}} \qquad \ldots \quad \ldots \quad \ldots \quad \ldots (\text{٥} - \text{٨})$$

مثال:

لحساب قيمة الانحراف المعيارى من مجموعة أو عدة مجموعات من البيانات (مثل حساب تباين أعمار ستة أشخاص للثلاث عينات السابق الإشارة إليها) نأخذ الجذر التربيعى لقيم تباين كل مجموعة. وإذا استخدمنا نفس بيانات الجدول رقم (٥-٣) فإن الانحراف المعيارى لكل عينة يكون كالآتى:

العينة الأولى = $\sqrt{\text{١}}$ = ١ سنة

العينة الثانية = $\sqrt{\text{١,٦٧}}$ = ١,٢٩ سنة

العينة الثالثة = $\sqrt{\text{٣}}$ = ١,٧٣ سنة.

وبالمثل يحسب الانحراف المعيارى للعينة كتقدير غير متحيز للانحراف المعيارى

للمجتمع الذى أخذت منه، أو ما يسمى أيضاً «أحسن تقدير للانحراف المعيارى للمجتمع ($\hat{ع}$)» – وللحصول على هذا التقدير غير المتحيز للانحراف المعيارى للمجتمع فإن الانحراف المعيارى للعينة يصحح بعامل «بسل» وذلك على النحو التالى:

$$\hat{ع} = \text{الانحراف المعيارى للعينة} \times \sqrt{\frac{ن}{ن - ١}}$$

$$= \sqrt{\frac{\text{مجـ}(س - \bar{س})^٢}{ن}} \times \sqrt{\frac{ن}{ن - ١}} \quad ... \quad ... \quad (٥ - ٩)$$

وللتبسيط والاختصار توضع الصيغة (٥ – ٩) على النحو التالى:

$$ع = \sqrt{\frac{\text{مجـ}(س - \bar{س})^٢}{ن}} \quad ... \quad ... \quad (٥ - ١٠)$$

ومما تجدر الإشارة إليه أن التباين والانحراف المعيارى – كمقياسين كميين للاختلاف أو التبعثر (للتشتت) – يختلفان فى مجال الاستفادة منهما فى التحليلات الجغرافية، فالمقياس الأول (التباين) يصلح لأغراض التنبؤ والتكهن ويستعمل بكثرة فى الاختبارات الاحصائية، بينما يصلح الانحراف المعيارى للأغراض الوصفية فقط حيث أنه يقاس بنفس وحدات الصفة المتغيرة.

أمثلة نوعية:

بعد ذلك العرض الموجز لطريقة حساب كل من التباين والانحراف المعيارى نحاول الآن تطبيقهما على بيانات نوعية فى مجال الدراسات الجغرافية حتى نتعرف على مجال الاستفادة منهما فى التحليل الكمى الجغرافى. فالجدول رقم (٥ –٤) يوضح كمية الأمطار السنوية التى سجلها مرصد Bidston – انجلترا (السابق الإشارة إليها) وذلك فى الفترة من ١٩٠١ إلى ١٩٣٠، وبحساب كل من الانحراف المتوسط، التباين والانحراف المعيارى بالطرق السابق ذكرها (المعادلات ٥

- ٨) وجد أن قيمة كل من الانحراف المتوسط والانحراف المعيارى هى على الترتيب ٢,٧٩ بوصة و ٣,٤٥ بوصة. ومن الواضح أنه إذا قارنا الانحراف المتوسط (٢,٧٩ بوصة) وقيمة الانحراف المعيارى (٣,٤٥ بوصة) وهما يقيسان التبعثر والتشتت لنفس المجموعة من كميات المطر - نجد أن قيمتهما مختلفتان وهذا أمر متوقع لأنه على الرغم من أنهما يقيسان نفس درجة الإختلاف بين قيم مجموعة واحدة من البيانات إلا أن كلاً منهما يحسب بطريقة مختلفة عن الآخر. كذلك فإنه يمكن استنتاج علاقة اعتبارية تربط بين الانحراف المعيارى والانحراف المتوسط للتوزيعات المتوسطة الالتواء والتى يعبر عنها بالصيغة: الانحراف المتوسط = $\frac{٤}{٥}$× الانحراف المعيارى، أو بمعنى آخر أن الانحراف المعيارى أكبر من الانحراف المتوسط بنسبة ٢٥٪، أى أن الانحراف المعيارى = ١,٢٥ × الانحراف المتوسط، ويظهر ذلك جلياً من النتاذج التى حصلنا منه الجدول رقم (٥ -٤) والتى توضح أن العامل الذى يرجح به الانحراف المتوسط يجب أن يكون ١,٢٤ ليعطى نفس الانحراف المعيارى الناتج بالمعادلة (٥ -٨).

جدول رقم (٥ -٤): حساب الانحراف المتوسط والانحراف المعيارى لكمية المطر السنوى فى مرصد Bidston - انجلترا فى الفترة من ١٩٠١ - ١٩٣٠

السنة	القيم	ح	ح٢
١٩٠١	٢٥,١٩	-٣,٢٦	١٠,٦
١٩٠٢	٢٥,٥٧	-٢,٨٨	٨,٣
١٩٠٣	٣٤,٤٢	+٥,٩٧	٣٥,٦
١٩٠٤	٢٥,١٨	-٣,٢٧	١٠,٧
١٩٠٥	٢٤,٠١	-٤,٤٤	١٩,٦
١٩٠٦	٢٨,٠٨	-٠,٣٧	٠,١
١٩٠٧	٢٦,٥٧	-١,٨٨	٣,٥
١٩٠٨	٢٨,٩٠	+٠,٤٥	٠,٢
١٩٠٩	٢٨,٤٥	صفر	صفر

١٩١٠	٢٨,٥٩	,١٤+	٠,٠٢
١٩١١	٢٥,٢٧	٣,١٨-	١٠,١
١٩١٢	٣٠,١٧	١,٧٢+	٢,٩٥
١٩١٣	٢٥,٧٨	٢,٦٧-	٧,١
١٩١٤	٢٦,٠٢	٢,٤٣-	٥,٩
١٩١٥	٢٦,٨٣	١,٦٢-	٢,٦
١٩١٦	٢٤,٨٧	٣,٥٨-	١٢,٨
١٩١٧	٣٠,٥٩	٢,١٤+	٤,٦
١٩١٨	٣١,٩٣	٣,٤٨+	١٢,١
١٩١٩	٢٩,١٢	٠,٦٧+	٠,٤٥
١٩٢٠	٣٣,٤٣	٤,٨٩+	٢٣,٩
١٩٢١	٢٢,٤٧	٥,٩٨-	٣٥,٧
١٩٢٢	٢٥,٩٧	٢,٤٨-	٦,١٥
١٩٢٣	٣٠,٩٢	٢,٤٧+	٦,١
١٩٢٤	٣٢,٨٧	٤,٤٢+	١٩,٦
١٩٢٥	٢٨,٠٠	٠,٤٥-	٠,٢
١٩٢٦	٢٨,٩٥	٠,٥٠+	٠,٢٥
١٩٢٧	٣٤,٨١	٦,٣٦+	٤٠,٥
١٩٢٨	٢٩,١١	٠,٦٦+	٠,٤
١٩٢٩	٢٥,١٥	٣,٣٠-	١٠,٩
١٩٣٠	٣٥,٥٠	٨,٠٥+	٦٥,٠
المجموع	٨٥٣,٦٣	٨٣,٧١	٣٥٥,٩٢
العدد	٣٠	٣٠	٣٠
المتوسط = ٢٨,٤٥		الانحراف المتوسط = ٢,٧	التبايـــــن = ١١,٨٦ الانحراف المعيارى = ٣,٤٥

وبعد إجراء عمليات حسابية بسيطة باستخدام بيانات إنتاج خام الحديد فى أربع دول من دول غرب أوربا: بلجيكا، فرنسا، لوكسمبرج والمملكة المتحدة (جدول رقم ٤ – ١٢) يتم إستخراج قيم كل من المتوسط الحسابى والانحراف المتوسط والانحراف المعيارى ونسبة الانحراف المعيارى إلى الانحراف المتوسط لإنتاج كل دولة كما هو مبين فى الجدول رقم (٥ – ٥) التالى:

جدول (٥ – ٥)

الانحراف المتوسط والانحراف المعيارى لإنتاج الحديد الخام فى بلجيكا، فرنسا، لوكسمبرج والمملكة المتحدة (١٩٣٨ – ١٩٥٧)

الدولــــة	المتوسط الحسابى (آلاف الأطنان)	الانحراف المتوسط (آلاف الأطنان)	الانحراف المعيارى (آلاف الأطنان)	الانحراف المعيارى / الانحراف المتوسط
بلجيكا	٣٤,٤٥	١٣,٥	١٥,٥٥	١,١٨
فرنسا	٩٣١٠,٢	٤٢٨٣,٢	٤٩٦٠,٠	١,١٢
لوكسمبورج	١٤٤٨,٨	٤٣٧,٣	٥٢٧,٢	١,٢٠
المملكة المتحدة	٤٤١٨,٩٥	٤٨٠,١	٦٥٦,٥	١,٣٧

ويلاحظ من الجدول السابق أن الانحراف المعيارى، فى كل الحالات، أكبر من الإنحراف المتوسط، كما أن النسبة بينهما تتراوح بين ١,١٦ و ١,٣٧. ويرجع السبب فى ذلك إلى أنه أثناء عملية تربيع قيم المفردات ثم أخذ الجذر التربيعى لمجموع مربعات هذه القيم، فإن الانحرافات الكبيرة تحمل وزناً متزايداً، بينما تحمل الانحرافات الصغيرة وزناً قليلاً، ومعنى ذلك أن الانحراف المعيارى يظل دائماً أكبر من الانحراف المتوسط. وتتوقف درجة الاختلاف بين الانحرافين على التكرار النسبى للقيم ومدى الانحراف الذى يحدده أكبر وأصغر الانحرافات الفردية. وبالإضافة إلى ذلك فإن هذه الخاصية للانحراف المعيارى تعكس حقيقة أن بيانات إنتاج الحديد فى دول غرب أوربا (جدول رقم ٤ – ١٢) تتصف بأنها غير متماثلة

فى توزيعها (ملتوية)، بينما تتصف بيانات كمية المطر السنوى فى مرصد Bidston - انجلترا بأنها تتوزع توزيعاً متماثلاً (أو معتدلا) Normally distributed تقريباً.

طرق أخرى لحساب التباين والانحراف المعيارى:

سبق أن لاحظنا أن عملية إستخراج التباين والانحراف المعيارى لمجموعة من البيانات (معادلات رقم: ٥ - ٤، ٥ - ٥، ٤ - ٧، ٥ - ٨) تكتنفها بعض الصعوبات الحسابية، ولذلك إذا كانت هناك عمليات أخرى لاختصار طرق حسابهما فقد يساعد ذلك على سهولة وتبسيط طرق استخراجهما من البيانات المتاحة. وهناك طريقان يمكن بهما تبسيط عمليات استخراج التباين والانحراف المعيارى: الطريقة الأولى تعتمد على إجراء تعديل جبرى فى معادلة التباين وتعرف باسم طريقة الفرق بين المربعين، والطريقة الأخرى هدفها تقليل عدد الحسابات وخطواتها وتعرف باسم طريقة التربيع أو طريقة الماكينة. وسنشرح فيما يلى كل طريقة على حدة.

طريقة الفرق بين المربعين:

سبق أن ذكرنا أن التباين هو متوسط مربع انحرافات قيم المفردات عن متوسطها الحسابى أو ما يعرف «بالانحراف التربيعى المتوسط» ومعادلته هى:

$$(\text{ع}^{2}) = \frac{\text{مجـ}\,(\text{س} - \bar{\text{س}})^{2}}{\text{ن}}$$

ويظهر أن العنصر الرئيسى فى تلك المعادلة هو $(\text{س} - \bar{\text{س}})^{2}$ والذى يطلق عليه جبرياً «الفرق بين المربعين»، ويمكن أن يكتب ذلك تفصيلياً بالصورة التالية:

$$(\text{س} - \bar{\text{س}})^{2} = (\text{س} - \bar{\text{س}})\,(\text{س} - \bar{\text{س}})$$

$$= \text{س}^{2} - 2\bar{\text{س}}\,\text{س} + \bar{\text{س}}^{2}$$

وبقسمة مفكوك القوس على عدد المفردات (ن) يمكن إعادة كتابة معادلة التباين على النحو التباين:

$$\text{التباين (ع}^{2}\text{)} = \frac{\text{مجـ س}^{2}}{\text{ن}} - \frac{\text{مجـ ٢ س}}{\text{ن}} \times \frac{\text{مجـ س}}{\text{ن}} + \frac{\text{مجـ }\bar{\text{س}}^{2}}{\text{ن}} \quad \dots (٥ - ١١)$$

وبالرغم من أنه يمكن إختصار كل جزء بمفرده من هذه الصورة المطولة لمعادلة التباين إلا أن عملية استخراج التباين بهذه المعادلة تبدو أكثر تعقيداً، ولذا يمكن تبسيط المكونات المفردة لهذه المعادلة عن طريق إستبدالها بمكونات أخرى.

فقد سبق القول أن سَ (المتوسط الحسابى) $= \frac{\text{مجـ س}}{\text{ن}}$ ، وعلى ذلك فإنه يمكن وضع (سَ) بدلاً من $\left(\frac{\text{مجـ س}}{\text{ن}}\right)$ وطالما أن قيمة (سَ) ثابتة فإنها ستظل كذلك فى بقية المعادلة، وبناء على ذلك إذا أضفنا قيمة (سَ) بعدد (ن) من المرات ثم قسمت بواسطة (ن) فإن الإجابة ستكون سَ، أى أن :

$$\frac{\text{مجـ س}}{\text{ن}} = \bar{\text{س}}$$

ونحاول الآن تبسيط معادلة التباين (٥ − ١١) باستبدال سَ لكل من

$$\frac{\text{مجـ س}}{\text{ن}} \text{ ، } \frac{\text{مجـ }\bar{\text{س}}}{\text{ن}} \text{ .}$$

$$\text{ع}^{2} = \frac{\text{مجـ س}^{2}}{\text{ن}} - ٢\bar{\text{س}} \times \text{س} + \bar{\text{س}}^{2}$$

$$= \frac{\text{مجـ س}^{2}}{\text{ن}} - ٢\bar{\text{س}}^{2} + \bar{\text{س}}^{2}$$

$$= \frac{\text{مجـ س}^{2}}{\text{ن}} - \bar{\text{س}}^{2} \quad \dots \quad \dots \quad \dots \quad (٥ - ١٣)$$

وبما أن الانحراف المعيارى هو الجذر التربيعى للتباين فإن صيغة الانحراف المعيارى تكتب كما يلى:

$$\text{الانحراف المعيارى (ع)} = \sqrt{\frac{\text{مجـ س}^{2}}{\text{ن}} - \bar{\text{س}}^{2}} \quad \dots \quad \dots \quad (٥ - ١٣)$$

ويلاحظ من المعادلتين (٥ −١٢، ٥− ١٣) أنهما لايتطلبان إلا حسابات قليلة، إذ يتم تربيع قيمة كل مفردة على حدة ثم تجمع مربعات هذه القيم وتقسم

على عدد المفردات ثم نطرح منها مربع المتوسط الحسابى للقيم فنحصل على التباين. وللحصول على الانحراف المعيارى نستخرج الجذر التربيعى لقيمة التباين كما ذكرنا آنفاً.

طريقة التربيع (طريقة الماكينة):

يمكن أيضاً حساب مجموع مربع الانحرافات عن المتوسط الحسابى عند حساب التباين لمجموعة من البيانات باتباع الخطوات التالية:

مجموع مربع الانحرافات عن المتوسط الحسابى = $\text{مجـ}(\text{س} - \bar{\text{س}})^{٢}$، وهو المكون الرئيسى لمعادلة التباين (معادلة رقم ٥ − ٥) ويساوى:

$$\text{مجـ}(\text{س} - \bar{\text{س}})^{٢} = [(\text{س}_{١} - \bar{\text{س}})^{٢} + (\text{س}_{٢} - \bar{\text{س}})^{٢} + .. + (\text{س}_{\text{ن}} - \bar{\text{س}})^{٢}]$$

$$= [(\text{س}_{١}^{٢} - ٢\text{س}_{١}\bar{\text{س}} + \bar{\text{س}}^{٢}) + (\text{س}_{٢}^{٢} - ٢\text{س}_{٢}\bar{\text{س}} + \bar{\text{س}}^{٢}) + ... + (\text{س}_{\text{ن}}^{٢} - ٢\text{س}_{\text{ن}}\bar{\text{س}} + \bar{\text{س}}^{٢})]$$

$$= \Big[(\text{س}_{١}^{٢} + \text{س}_{٢}^{٢} + ... + \text{س}_{\text{ن}}^{٢}) - ٢\bar{\text{س}}(\text{س}_{١} + \text{س}_{٢} + ... + \text{س}_{\text{ن}}) + (\bar{\text{س}}^{٢} + \bar{\text{س}}^{٢} + ... + \bar{\text{س}}^{٢})\Big]$$

$$= \text{مجـ س}^{٢} - ٢\bar{\text{س}}\ \text{مجـ س} + \text{ن}\,\bar{\text{س}}^{٢} \quad ... \quad ... \quad ... \quad ...(٥ - ١٤)$$

وبالتعويض عن قيمة $\bar{\text{س}}$ التى تساوى $\frac{\text{مجـ س}}{\text{ن}}$ فى المعادلة السابقة:

$$= \text{مجـ س}^{٢} - ٢\,\frac{\text{مجـ س} \times \text{مجـ س}}{\text{ن}} + \text{ن}\,\frac{\text{مجـ س} \times \text{مجـ س}}{\text{ن} \times \text{ن}}$$

$$= \text{مجـ س}^{٢} - \frac{\text{مجـ س} \times \text{مجـ س}}{\text{ن}}$$

$$= \text{مجـ س}^2 - \frac{(\text{مجـ س})^2}{\text{ن}} \quad ... \quad ... \quad ... \quad ... \quad (٥ - ١٥)$$

ويطلق على القيمة $\frac{(\text{مجـ س})^2}{\text{ن}}$ بعامل التصحيح، أى أن مجموع مربع الانحرافات عن المتوسط الحسابى يساوى مجموع مربع كل قيمة من قيم المفردات مطروحاً منه عامل التصحيح وهو عبارة عن مربع مجموع القيم مقسوماً على عددها. وعلى ذلك فإن صيغة كل من التباين والانحراف المعيارى ستكون حينئذ:

$$\text{التباين } (\text{ع}^2) = \frac{1}{\text{ن}} \left[\text{مجـ س}^2 - \frac{(\text{مجـ س})^2}{\text{ن}} \right] \quad ... (٥ - ١٦)$$

$$\text{الانحراف المعيارى} = \sqrt{\frac{1}{\text{ن}} \left(\text{مجـ س}^2 - \frac{(\text{مجـ س})^2}{\text{ن}} \right)} \quad ... (٥ - ١٧)$$

ويلاحظ أيضاً أن تطبيق أى من المعادلتين يتطلب حسابات أقل من طريقة الفرق بين المربعين السابق وشرحها، ذلك لأن كل قيمة مفردة يتم تربيعها وتجمع مربعات هذه القيم ثم تجمع قيم المفردات ويتم تربيع هذا المجموع ويقسم على عدد المفردات ويطرح الناتج من مجموع مربعات القيم ثم يقسم الناتج الأخير على عدد المفردات فنحصل على التباين، وبأخذ الجذر التربيعى له نحصل على الانحراف المعيارى.

ولتوضيح تطبيق الطرق الثلاث السابقة لحساب التباين والانحراف المعيارى نستعين بالمثال التالى:

مثال: فى دراسة لبيان مجال نفوذ مدينة ما، إتخذت خدمات السكك الحديدية بين تلك المدينة والمدن المجاورة معياراً لذلك. ولنفترض أن هناك ٢٥ مدينة تتم خدمتها بالقطارات، وأن عدد القطارات اليومية لكل من هذه المدن يختلف كما هو مبين فى العمود الثانى من الجدول رقم (٥ - ٦)، وبعملية حسابية بسيطة يمكن

إيجاد متوسط عدد القطارات ى اليوم الواحد بين المدينة، موضع الدراسة – وبين أى مدينة مجاورة، ووجد أنه ٩,٦ قطاراً فى اليوم. ومن الواضح أن هناك مدناً يزيد عدد القطارات التى تصل إليها عن هذا المتوسط زيادة كبيرة وأخرى يقل عدد القطارات عنه قلة واضحة، أى أنها بيانات تمثل متغيراً غير متصلاً Disctete Variable، ولذا فمن المفيد معرفة إنتشار (تشتت) قيم المفردات حول هذا المتوسط، وأدق المقاييس لمعرفة ذلك هو الانحراف المعيارى.

ويبين الجدول رقم (٥- ٦) طرق حساب الانحراف المعيارى عن طريق حساب التباين أولاً بالمعادلات التى سبق ذكرها من قبل وهى:

$$\text{التباين} = \frac{\text{مجـ}(\text{س} - \bar{\text{س}})^{2}}{\text{ن}} \qquad \text{(الطريقة الأولى)}$$

$$\text{التباين} = \frac{\text{مجـ س}^{2}}{\text{ن}} - \bar{\text{س}}^{2} \qquad \text{(الطريقة الثانية)}$$

$$\text{التباين} = \frac{1}{\text{ن}} \left[\text{مجـ س}^{2} - \frac{(\text{مجـ س})^{2}}{\text{ن}} \right] \qquad \text{(الطريقة الثالثة)}$$

وكما هو موضح تعطى كل طريقة نفس قيمة التباين أى ٢٦,٠، وبذلك يكون الانحراف المعيارى فى كل ٥,١ قطاراً.

جدول رقم (٦٥)
حساب الانحراف المعيارى لعدد القطارات اليومية بين مدينة مركزية ومدن مجاورة باستخدام ثلاث طرق مختلفة

المدن المجاورة	عدد القطارات يومياً (س)	الطريقة الأولى		الطريقة الثانية والثالثة
		الانحراف (س - سَ)	الانحراف (س - سَ)٢	مربع القيم (س٢)
١	١	-٨,٦	٧٣,٩٦	١
٢	٢	-٧,٦	٥٧,٧٦	٤
٣	٣	-٦,٦	٤٣,٥٦	٩
٤	٣	-٦,٦	٤٣,٥٦	٩
٥	٤	-٥,٦	٣١,٣٦	١٦
٦	٥	-٤,٦	٢١,١٦	٢٥
٧	٦	-٣,٦	١٢,٩٦	٣٦
٨	٦	-٣,٦	١٢,٩٦	٣٦
٩	٨	-١,٦	٢,٥٦	٦٤
١٠	٨	-١,٦	٢,٥٦	٦٤
١١	٨	-١,٦	٢,٥٦	٦٤
١٢	١٠	+٠,٤	٠,١٦	١٠٠
١٣	١٠	+٠,٤	٠,١٦	١٠٠
١٤	١٠	+٠,٤	٠,١٦	١٠٠
١٥	١٠	+٠,٤	٠,١٦	١٠٠
١٦	١١	+١,٤	١,٩٦	١٢١
١٧	١١	+١,٤	١,٩٦	١٢١
١٨	١٢	+٢,٤	٥,٧٦	١٤٤
١٩	١٢	+٢,٤	٥,٧٦	١٤٤
٢٠	١٤	+٤,٤	١٩,٣٦	١٩٦
٢١	١٥	+٥,٤	٢٩,١٦	٢٢٥
٢٢	١٥	+٥,٤	٢٩,١٦	٢٢٥
٢٣	١٧	+٧,٤	٥٤,٧٦	٢٨٩
٢٤	١٩	+٩,٤	٨٨,٣٦	٣٦١
٢٥	٢٩	+١٠,٤	١٠٨,١٦	٤٠٠
المجموع	٢٤٠		٦٥٠,٠٠	٢٩٥٤

$$\therefore \text{المتوسط الحسابى } (\bar{\text{س}}) = \frac{٢٤٠}{٢٥} = ٩,٦$$

الطريقة الأولى:

$$\text{التباين} = \frac{٦٥٠}{٢٥} = ٢٦,٠$$

الطريقة الثانية:

$$\text{التباين} = \frac{٢٩٥٤}{٢٥} - (٩,٦)^{٢}$$

$$= ١١٨,١٦ - ٩٢,١٦ = ٢٦,٠$$

الطريقة الثالثة:

$$\text{التباين} = \frac{١}{٢٥}\left(٢٩٥٤ - \frac{(٢٤٠)^{٢}}{٢٥}\right)$$

$$= \frac{١}{٢٥} \times ٦٥٠ = ٢٦,٠$$

والانحراف المعيارى فى كل الحالات = $\sqrt{٢٦}$ = ٥,١ قطاراً

ايجاد التباين والانحراف المعيارى للبيانات المبوبة:

هناك عدة طرق لحساب التباين والانحراف المعيارى للبيانات المبوبة فى جداول التوزيعات التكرارية ولكنها لاتختلف عن طريقة حسابهما للبيانات غير المبوبة. ويحسب التباين بالصيغة التالية:

$$\text{التباين } (\text{ع}^{٢}) = \frac{\text{مجـ ك } (\text{س} - \bar{\text{س}})^{٢}}{\text{مجـ ك}} \quad \ldots \quad \ldots \quad \ldots \quad \ldots (٤ - ١٨)$$

حيث مجـ ك = مجموع التكرارات، م هى مراكز الفئات ولصعوبة استخدام هذه الصيغة يمكن وضعها على الصورة التالية:

$$(\text{ع_}^{٢}) = \frac{\text{مجـ م}^{٢}\text{ك}}{\text{مجـ ك}} - \left(\frac{\text{مجـ م ك}}{\text{مجـ ك}}\right)^{٢} \quad .. \quad ... (٥ - ١٩)$$

ويكون الانحراف المعيارى:

$$(\text{ع_}) = \sqrt{\frac{\text{مجـ م}^{٢}\text{ك}}{\text{مجـ ك}} - \left(\frac{\text{مجـ م ك}}{\text{مجـ ك}}\right)^{٢}} \quad ... \quad ... (٥ - ٢٠)$$

وتسمى هذه الصيغة لحساب الانحراف المعيارى بالطريقة المطولة.

ويبين الجدول رقم (٥ – ٧) خطوات حساب التباين والانحراف المعيارى من جدول التوزيع التكرارى لعدد القطارات اليومية بين مدينة مركزية ومدن مجاورة الذى تم إعداده على أساس تبويب البيانات الموضحة فى الجدول رقم (٥–٦) فى فئات كما هو متبع فى إعداد الجداول التكرارية، وذلك مع إبقاء المدى المطلق Range للقيم فى أى فئة صغيراً حتى يقل مدى الخطأ الناتج عن التعميم، ومع عدم تجاوز عدد الفئات عن عشرة فئات.

جدول رقم (٥ – ٧)

حساب التباين والانحراف المعيارى من البيانات المبوبة لعدد القطارات اليومية بين مدينة مركزية ومدن مجاورة المبين فى الجدول رقم (٥ –٦)

الفئات	الحدود الحقيقية للفئات	مراكز الفئات (م)	التكرار (ك)	م × ك	م٢ × ك
١ – ٢	٠,٥ – ٢,٥	١,٥	٢	٣,٠	٤,٥٠
٣ – ٤	٢,٥ – ٤,٥	٣,٥	٣	١٠,٥	٣٦,٧٥
٥ – ٦	٤,٥ – ٦,٥	٥,٥	٣	١٦,٥	٩٠,٧٥
٧ – ٨	٦,٥ – ٨,٥	٧,٥	٣	٢٢,٥	١٦٨,٧٥
٩ – ١٠	٨,٥ – ١٠,٥	٩,٥	٤	٣٨,٠	٣٦١,٠٠
١١ – ١٢	١٠,٥ – ١٢,٥	١١,٥	٤	٤٦,٠	٥٢٩,٠٠
١٣ – ١٤	١٢,٥ – ١٤,٥	١٣,٥	١	١٣,٥	١٨٢,٢٥
١٥ – ١٦	١٤,٥ – ١٦,٥	١٥,٥	٢	٣١,٠	٤٨٠,٥٠
١٧ – ١٨	١٦,٥ – ١٨,٥	١٧,٥	١	١٧,٥	٣٠٦,٢٥
١٩ – ٢١	١٨,٥ – ٢٠,٥	١٩,٥	٢	٣٩	٧٦٠,٥٠
			٢٥	٢٣٧,٥	٢٩٢٠,٢٥

ولحساب التباين والانحراف المعيارى للبيانات السابقة نجد أن:

$$\text{التباين (ع}^2\text{)} = \frac{٢٩٢٠,٢٥}{٢٥} - \left(\frac{٢٣٧,٥}{٢٥}\right)^{٢}$$

$$= ١١٦,٨١ - ٩٠,٢٥ = ٢٦,٥٦$$

والانحراف المعيارى (ع) = $\sqrt{٢٦,٥٦}$ = ٥,١٥٤ قطاراً

وواضح أن القيمة التى حصلنا عليها للانحراف المعيارى (٥,١٥٤) تختلف عن النتيجة السابقة (٥,١) من البيانات التفصيلية. وكما ذكرنا سابقاً عند حساب

المتوسط الحسابى من بيانات مبوبة، فإن الخطأ الناشئ عن العمليات الحسابية لإيجاد الانحراف المعيارى للبيانات المبوبة يعتبر ضئيلاً على الرغم من العدد المحدود الفئات التوزيع الموضحة بالجدول، وهذا الخطأ من وجهة النظر الاحصائية يعد مقبولاً خصوصاً إذا عرفنا ما توفره طريقة تجميع البيانات وتبويبها من وقت فى إجراء العمليات الحسابية عنه مع الطريقة العادية للبيانات التفصيلية التى تتضمن عدداً كبيراً من قيم المفردات والتى غالباً ما تحتوى على قيم شاذة أو متطرفة.

ولتسهيل العمليات الحسابية لإيجاد التباين والانحراف المعيارى من بيانات جدول توزيع تكرارى منتظم (أى الذى تتساوى فيه أطوال الفئات) اتستخدم طريقة أخرى تعرف باسم **طريقة الانحرافات المختصرة**. وفى هذه الطريقة يختار وسط فرضى من بين مراكز الفئات ثم نحسب الانحرافات عن هذا الوسط الفرضى بعد إختصارها وذلك بقسمتها على طول الفئة ثم نربعها ونوجد متوسطها بعد ذلك، وهذا يعنى أننا طرحنا مقداراً ثابتاً من مراكز الفئات فيكون الانحراف المعيارى للانحرافات (أى القيم بعد طرح الوسط الفرضى منها) هو نفسه الانحراف المعيارى لمراكز الفئات الأصلية. وفى هذه الحالة يمكن حساب التباين والانحراف المعيارى بإستخدام الصيغة الآتية:

$$\text{التباين } (\text{ع}^{2}) = \text{ل} \times \left[\frac{\text{مجـ ح}'^{2}\text{ ك}}{\text{مجـ ك}} - \left(\frac{\text{مجـ ح}'\text{ ك}}{\text{مجـ ك}} \right)^{2} \right] \quad \ldots (٥ - ٢١)$$

$$\text{الانحراف المعيارى } (\text{ع}) = \text{ل} \times \sqrt{ \frac{\text{مجـ ح}'^{2}\text{ ك}}{\text{مجـ ك}} - \left(\frac{\text{مجـ ح}'\text{ ك}}{\text{مجـ ك}} \right)^{2} } \quad \ldots \; \ldots (٥ - ٢٢)$$

حيث ل = طول الفئة، ح َ = الانحراف المختصر عن الوسط الفرضى ويوضح الجدول رقم (٥ – ٨) خطوات إستخراج الانحراف المعيارى من جدول التوزيع التكرارى لأطوال مائة رافد نهرى (بالأمتار) المبينة بالجدول رقم (٢ – ١١).

جدول رقم (٥-٨)

حساب الانحراف المعيارى لأطوال مائة رافد نهرى (مترى لأحد الأحواض النهرية فى إحدى المناطق باستخدام طريقة الانحرافات المختصرة

الفئات (متر)	الحدود الحقيقية للفئات	مراكز الفئات (س)	التكرار ك	الانحرافات (ح)	الانحرافات مختصرة (حَ)	(حَ × ك)	(حَ^{٢} × ك)	ك (حَ × ١)^{٢}
٥٥٠ – ٦٤٠	٥٤٥ – ٦٤٥	٥٩٥	١	-٤٠٠	-٤	-٤	١٦٠	٩
٦٥٠ – ٧٤٠	٦٤٥ – ٧٤٥	٦٩٥	٢	-٣٠٠	-٣	-٦	١٨	٨
٧٥٠ – ٨٤٠	٧٤٥ – ٨٤٥	٧٩٥	٩	-٢٠٠	-٢	-١٨	٣٦	٩
٨٥٠ – ٩٤٠	٨٤٥ – ٩٤٥	٨٩٥	٢٢	-١٠٠	-١	-٢٢	٢٢	صفر
٩٥٠ – ١٠٤٠	٩٤٥ – ١٠٤٥	٩٩٥	٣٣	صفر	صفر	صفر	صفر	٣٣
١٠٥٠ – ١١٤٠	١٠٤٥ – ١١٤٥	١٠٩٥	٢٢	١٠٠	١	٢٢	٢٢	٨٨
١١٥٠ – ١٢٤٠	١١٤٥ – ١٢٤٥	١١٩٥	٨	٢٠٠	٢	١٦	٣٢	٧٢
١٢٥٠ – ١٣٤٠	١٢٤٥ – ١٣٤٥	١٢٩٥	٢	٣٠٠	٣	٦	١٨	٣٢
١٣٥٠ – ١٤٤٠	١٣٤٥ – ١٤٤٥	١٣٩٥	١	٤٠٠	٤	٤	١٦	٢٥
المجموع = مجـ			١٠٠			-٢	١٨٠	٢٧٦

ولإيجاد الإنحراف المعيارى باستخدام بيانات الجدول السابق نجد أن

$$\text{الانحراف المعيارى (ع)} = ١٠٠ \times \sqrt{\left[\frac{١٨٠}{١٠٠} - \left(\frac{-٢}{١٠٠}\right)^{٢}\right]}$$

$$= ١٠٠ \times \sqrt{١,٧٩٩٦}$$

= ١٣٤,١٤ متراً.

وتجدر الإشارة هناك إلى أنه يجب أن يسبق عملية استخراج الإنحراف المعيارى للبيانات المبوبة إجراء تحقيق وفحص صحة الحسابات فى عملية الجدولة حتى نتفادى الوقوع فى خطأ، وحتى يمكن عمل التصحيحات اللازمة قبل إيجاد قيمة هذا المقياس. وفى هذا الصدد يمكن تطبيق طريقة «اختبار تشارلبير Charlier's Test» لمراجعة حساب الانحراف المعيارى باستخدام المتطابقات:

$$\text{مجـ ك } (حَ + ١)^{٢} = \text{مجـ ك } حَ^{٢} + ٢ \text{ مجـ ك } حَ + \text{مجـ ك}$$

$$٢٧٦ = ١٨٠ + ٢ (-٢) + ١٠٠$$

$$٢٧٦ = ٢٧٦$$

ويلاحظ أن طرفى المعادلة متساويان مما يدل على أن العمليات الحسابية فى الجدولة قد تمت بدقة، وأن أى خطأ صغير يبقى بين الأرقام يمكن اعتباره خطأ ناتج عن التعميم والناتج بدوره عن اتساع الفئات التكرارية.

تصحيح التباين والانحراف المعيارى:

عند حساب كل من التباين والانحراف المعيارى للبيانات المبوبة فإنه يكون معرضاً لبعض الخطأ الناتج عن تجميع البيانات فى فئات (أخطاء التجميع)، بمعنى أننا نفترض أن مفردات كل فئة تتجمع أو تمثلها قيمة واحدة هى مركز الفئة وهو افتراض تقريبى الغرض منه تسهيل العمليات الحسابية. وبما أن القيم الحقيقية للمفردات تختلف داخل الفئة الواحدة فإن الفرق بين القيم الحقيقية والقيم

المفروضة (مراكز الفئات) تؤدى إلى بعض الأخطاء فى نتائج العمليات الحسابية، وبالتالى فإن نتائج المقاييس الإحصائية، ومن بينها مقاييس الانحراف المحسوبة من جدول توزيع تكرارى لاتنطبق تماماً على النتائج للمقاييس المناظرة المحسوبة من البيانات التفصيلية (راجع نتائج حساب كل من التباين والانحراف المعيارى لبيانات الجدولين رقم (٥ – ٦) (٥ – ٧). ولتعديل هذا النوع من الخطأ نستخدم الصيغة التالية:

$$\text{التباين المعدل (المصحح)} = \text{التباين المحسوب من البيانات المبوبة} - \frac{\text{ل}^{٢}}{١٢}$$

حيث ل هى طول الفئة، ومعامل التصحيح ($\frac{\text{ل}^{٢}}{١٢}$) المطروح يسمى تصحيح شبرد Sheppard's Correction، وبأخذ الجذر التربيعى لباقى الطرح من الصيغة السابقة ينتج الإنحراف المعيارى المصحح. ففى المثال الموضح فى الجدول رقم (٥ – ٧) وجدنا أن التباين قيمته ٢٦,٥٦ وهو يختلف عن التباين للبيانات التفصيلية (قبل التبويب) وقيمته ٢٦,٠، وعلى هذا يكون التباين المعدل هو.

$$\text{التباين المعدل} = ٢٦,٥٦ - \frac{(٢)^{٢}}{١٢}$$

$$= ٢٦,٥٦ - ٠,٣٣ = ٢٦,٢٣$$

ويكون الانحراف المعيارى المصحح $= \sqrt{٢٦,٢٣} = ٥,١٣$ قطاراً.

ومن الواضح أن تصحيح شبرد لايستخدم إلا فى حالة توزيعات المتغيرات المتصلة (المتغيرات المستمرة Coutiueous Variables) حيث «أطراف التوزيع» فيها تؤول تدريجياً إلى الصفر فى كلا الإتجاهين، أو بمعنى آخر لايسرى على التوزيعات النونية (على شكل حرف U) أو التوزيعات الشديدة الإلتواء. ويختلف الإحصائيون فى متى وما إذا كان تصحيح شبرد يجب تطبيقه لتعديل التباين والانحراف المعيارى المحسوبين من بيانات مبوبة. إلا أنه يجب عدم تطبيقه إلا بعد فحص دقيق للبيانات إذ أنه كثيراً ما يؤدى إلى مبالغة فى التصحيح وهذا يؤدى بدوره إلى استبدال الخطأ

القديم بخطأ جديد. وعلى العموم فإنه يفضل عدم إجراء هذا التصحيح للجداول التكرارية التى تقل فيها عدد التكرارات عن ١٠٠٠ أو تقل بها عدد الفئات عن ٢٠ فئة.

مؤشرات الاختلاف (التباين) Variability Indices

ذكرنا أنه لمقارنة مجموعتين من البابنات يجب استخدام متوسطيهما أولاً ثم تشتتيهما بعد ذلك بواسطة استخدام عدة مقاييس احصائية للمتوسطات والتشتت. وكما سبق أن لاحظنا على هذه المقاييس أنها تعطى فى النهاية قيماً مطلقة (غير نسبية) من نفس وحدات الصفة المتغيرة المستخرجة منها (مثل عدد الأمتار، البوصات الأطنان، عدد السكان، عدد الدرجات: مئوية أو فهرنيهتية بالنسبة لدرجة الحرارة مثلاً ... إلخ). وبما أنه من الطبيعى أن تختلف وحدات الصفة المتغيرة من مجموعة من القيم إلى مجموعة أخرى مما يؤدى بدوره إلى إختلاف وحدات مقاييس المتوسطات والتشتت، فإن المقارنة لاتصلح إلا بين المجموعات ذات النوع الواحد من الصفة المتغيرة للبيانات، فلا يعقل مثلاً أن نقارن بين كمية البضائع (بالأطنان) بأطوال السكك الحديدية (بالكيلو مترات) المنقولة عليها. وللتغلب على الاختلاف فى وحدات هذه المقاييس وحتى تكون المقارنة صحيحة ينبغى استخدام وحدات من نفس النوع أو استخدام أعداد مجردة خالية من الوحدات التى يمكن الحصول عليها بقسمة عددين من نوع واحد (أى من نفس الوحدات) أحدهما على الآخر. فمثلاً نقسم أحد مقاييس الإنحراف على أقرب مقاييس المتوسطات إليه فى طبيعته ويعبر عن ناتج القسمة كنسبة مئوية (أى نضرب الناتج فى ١٠٠٪)، وبذلك نكون قد حولنا قيم مقاييس الإنحراف من قيم مطلقة إلى قيم مئوية (بالنسبة للمتوسطات)، وهذه القيم المئوية هى التى نطلق عليها اسم «مؤشرات الاختلاف أو التباين». ونظراً للأهمية والفائدة الكبيرة لهذه المؤشرات عند عقد المقارنات بين المجموعات من البيانات فقد أصبحت تستخدم اليوم على نطاق كبير فى التحليل الكمى للجغرافية.

ومؤشرات الاختلاف أنواع منها مؤشر الاختلاف الذى يستخدم قيمة الوسيط والانحراف الربيعى وذلك على النحو التالى:

$$\frac{\text{الانحراف الربيعى}}{\text{الوسيط}} \times ١٠٠٪ \quad ... \quad ... \quad ...(٥ - ٢٣)$$

والمؤشر الثانى للاختلاف هو ما يطلق عليه «مؤشر الاختلاف النسبى» Relative Variability Index ويستخدم قيمة المتوسط الحسابى والإنحراف المتوسط ويمكن الحصول عليه من الصيغة الآتية:

$$\frac{\text{الانحراف المتوسط}}{\text{المتوسط الحسابى}} \times ١٠٠٪ \quad ... \quad ... \quad ...(٥ - ٢٤)$$

أما إذا استخدمنا الانحراف المعيارى والمتوسط الحسابى فإننا نحصل على مؤشر للاختلاف يعرف باسم معامل الاختلاف Coefficient of Variation ويستخرج كالآتى:

$$\frac{\text{الانحراف المعيارى}}{\text{المتوسط الحسابى}} \times ١٠٠٪ \quad ... \quad ... \quad (٥ - ٢٥)$$

ويعد معامل الاختلاف أكثر مؤشرات الاختلاف انتشاراً، ويستخدم فى الحالات التالية:

١ - تؤخذ قيمة معامل الاختلاف كمؤشر أو دليل لمدى دقة القياس أو فى عملية الحصول على البيانات إذا أنه لايمكن أن نقارن بين الانحرافات المعيارية مباشرة لاختلاف قيمة المتوسطات الحسابية لهذه العمليات. ولكن يمكن القول أنه كلما قلت قيمة معامل الاختلاف كلما زادت دقة القياس والعكس صحيح. ونظراً لأن معامل الاختلاف للقياسات الحقلية يختلف عن معامل الاختلاف

للقياسات المعملية فقد وضعت حدوداً لقيمة المعامل يجب أن لاتتعداها القياسات فى كل حالة وإلا كانت النتائج المستحصلة غير دقيقة. وتتراوح قيمة معامل الاختلاف للقياسات الحقلية ما بين ١٠ – ٢٠ ٪ بينما تقل للقياسات المعملية عن ١٠ ٪. وفى حالة القياسات الحقلية إذا قلت قيمة معامل الاختلاف للبيانات عن ١٠ ٪ فإنه يجب مراجعة قيم البيانات خوفاً من وقوع بعض الأخطاء فى العمليات الحسابية، أما إذا زادت قيمته عن ٢٠٪ فإنه يستدل من ذلك على أن عملية القياس قد تمت تحت ظروف غير ملائمة مما يترتب عليه استخلاص نتائج لايعول عليها كثيراً، ولذا يجب التأكد من شروط عملية القياس قبل إجرائها.

٢– يستخدم معامل الاختلاف لمقارنة الاختلافات بين توزيع عينتين لنفس الصفة المتغيرة ولكنها مقاسة بوحدات مختلفة، فمثلاً إذا كانت لدينا بيانات عن توزيع درجة الحرارة فى إحدى المناطق لفترتين مختلفتين وكان متوسط الحرارة للفترة الأولى هو ٢٣ درجة مئوية والانحراف المعيارى ٢°م، وكان متوسط درجة الحرارة للفترة الثانية هو ٨٦ درجة فهرنهيتية والانحراف المعيارى ٦°ف. ففى هذه الحالة لايمكننا أن نقارن بين الانحراف المعيارى للفترتين وذلك لاختلاف وحدات القياس ولكن يمكن التغلب على ذلك بمقارنة معاملى الاختلاف للفترتين.

فمعامل الاختلاف للفترة الأولى = $\frac{٢}{٢٢}$ × ١٠٠ = ٨,٧ ٪

ومعامل الاختلاف للفترة الثانية = $\frac{٦}{٨٦}$ × ١٠٠ = ٦,٩٧ ٪

أى أن الاختلاف فى درجة الحرارة للفترة الأولى أكبر منه للفترة الثانية.

٣– يستخدم معامل الانحراف أيضاً فى حالة مقارنة التوزيعات ذات الفروق الكبيرة

ومؤشرات الاختلاف أنواع منها مؤشر الاختلاف الذى يستخدم قيمة الوسيط والانحراف الربيعى وذلك على النحو التالى:

$$\frac{\text{الانحراف الربيعى}}{\text{الوسيط}} \times ١٠٠\% \quad ... \quad ...(٥ - ٢٣)$$

والمؤشر الثانى للاختلاف هو ما يطلق عليه «مؤشر الاختلاف النسبى» Relative Variability Index ويستخدم قيمة المتوسط الحسابى والإنحراف المتوسط ويمكن الحصول عليه من الصيغة الآتية:

$$\frac{\text{الانحراف المتوسط}}{\text{المتوسط الحسابى}} \times ١٠٠\% \quad ... \quad ...(٥ - ٢٤)$$

أما إذا استخدمنا الانحراف المعيارى والمتوسط الحسابى فإننا نحصل على مؤشر للاختلاف يعرف باسم معامل الاختلاف Coefficient of Variation ويستخرج كالآتى:

$$\frac{\text{الانحراف المعيارى}}{\text{المتوسط الحسابى}} \times ١٠٠\% \quad ... \quad ... \quad (٥ - ٢٥)$$

ويعد معامل الاختلاف أكثر مؤشرات الاختلاف انتشاراً، ويستخدم فى الحالات التالية:

١ - تؤخذ قيمة معامل الاختلاف كمؤشر أو دليل لمدى دقة القياس أو فى عملية الحصول على البيانات إذا أنه لايمكن أن نقارن بين الانحرافات المعيارية مباشرة لاختلاف قيمة المتوسطات الحسابية لهذه العمليات. ولكن يمكن القول أنه كلما قلت قيمة معامل الاختلاف كلما زادت دقة القياس والعكس صحيح. ونظراً لأن معامل الاختلاف للقياسات الحقلية يختلف عن معامل الاختلاف

للقياسات المعملية فقد وضعت حدوداً لقيمة المعامل يجب أن لاتتعداها القياسات فى كل حالة وإلا كانت النتائج المستحصلة غير دقيقة. وتتراوح قيمة معامل الاختلاف للقياسات الحقلية ما بين ١٠ - ٢٠ ٪ بينما تقل للقياسات المعملية عن ١٠ ٪. وفى حالة القياسات الحقلية إذا قلت قيمة معامل الاختلاف للبيانات عن ١٠ ٪ فإنه يجب مراجعة قيم البيانات خوفاً من وقوع بعض الأخطاء فى العمليات الحسابية، أما إذا زادت قيمته عن ٢٠ ٪ فإنه يستدل من ذلك على أن عملية القياس قد تمت تحت ظروف غير ملائمة مما يترتب عليه استخلاص نتائج لايعول عليها كثيراً، ولذا يجب التأكد من شروط عملية القياس قبل إجرائها.

٢- يستخدم معامل الاختلاف لمقارنة الاختلافات بين توزيع عينتين لنفس الصفة المتغيرة ولكنها مقاسة بوحدات مختلفة، فمثلاً إذا كانت لدينا بيانات عن توزيع درجة الحرارة فى إحدى المناطق لفترتين مختلفتين وكان متوسط الحرارة للفترة الأولى هو ٢٣ درجة مئوية والانحراف المعيارى ٢°م، وكان متوسط درجة الحرارة للفترة الثانية هو ٨٦ درجة فهرنهيتية والانحراف المعيارى ٦°ف. ففى هذه الحالة لايمكننا أن نقارن بين الانحراف المعيارى للفترتين وذلك لاختلاف وحدات القياس ولكن يمكن التغلب على ذلك بمقارنة معاملى الاختلاف للفترتين.

فمعامل الاختلاف للفترة الأولى = $\frac{٢}{٢٢}$ × ١٠٠ = ٨,٧ ٪

ومعامل الاختلاف للفترة الثانية = $\frac{٦}{٨٦}$ × ١٠٠ = ٦,٩٧ ٪

أى أن الاختلاف فى درجة الحرارة للفترة الأولى أكبر منه للفترة الثانية.

٣- يستخدم معامل الانحراف أيضاً فى حالة مقارنة التوزيعات ذات الفروق الكبيرة

بين متوسطاتها حتى ولو كانت بياناتها مقيسة بنفس وحدات القياس. مثال ذلك إذا كان المتوسط الحسابى والانحراف المعيارى للملكية الزراعية فى منطقتين هو: فى المنطقة الأولى المتوسط = ١٠ أفدنة والانحراف المعيارى ١,٥ فداناً، وفى المنطقة الثانية متوسط الملكية = ٥ أفدنة والانحراف المعيارى ١,٠ فداناً. ولأول وهلة يمكن أن نقول، اعتماداً على القيم المطلقة للتشتت، أن الملكية فى المنطقة الأولى أكثر اختلافاً عن المنطقة الثانية وذلك إذا أهملنا المتوسط الحسابى لكلا المنطقتين، ولكن ذلك لايكون صحيحاً نظراً لاختلاف المتوسط الحسابى فى كل منهما. أما إذا حسبنا الاختلافات بالنسبة للمتوسط الحسابى نجد أن:

$$\text{معامل الاختلاف للمنطقة الأولى} = \frac{١,٥}{١٠} \times ١٠٠ = ١٥\%$$

$$\text{ومعامل الاختلاف للمنطقة الثانية} = \frac{١}{٥} \times ١٠٠ = ٢٠\%$$

وتشير النتيجة السابقة للاختلاف النسبى مقيساً بمعامل الاختلاف للملكية فى المنطقتين أن المنطقة الثانية أكثر اختلافاً من المنطقة الأولى لو أدخلنا المتوسط الحسابى فى الحساب عند تقييم النتائج، وهذه النتيجة تخالف تماماً النتيجة التى نحصل عليها بالاعتماد على القيم المطلقة لاختلاف أو تشتت الملكية فى المنطقتين، أو بمعنى آخر بالاعتماد على الانحراف المعيارى فقط لمقارنة درجة الاختلاف بينهما.

وكمثال تطبيقى، تم إجراء العمليات الحسابية الخاصة بكل مؤشر من مؤشرات الاختلاف السابق ذكرها باستخدام بيانات إنتاج خام الحديد فى أربع دول من دول غرب أوربا (جدول رقم ٤ - ١٢) واستخرجت قيم هذه المؤشرات ووضعت بالجدول التالي:

جدول رقم (٥ - ٩)

مؤشرات الاختلاف لإنتاج خام الحديد في بلجيكا وفرنسا ولوكسمبورج والمملكة المتحدة في الفترة من ٣٨ - ١٩٥٧ (نسب مئوية)

الدولـــة	الانحراف المعيارى (ألف طن)	«مؤشر الاختلاف» الانحراف الربيعى / المتوسط × ١٠٠	«الاختلاف النسبى» الانحراف المتوسط / المتوسط × ١٠٠	«معامل الاختلاف» الانحراف المعيارى / المتوسط × ١٠٠
		٪	٪	٪
بلجيكا	١٥,٥٥	٤١,٣٠	٣٨,٢	٤٥,١
فرنسا	٤٩٦٠,٠	٤٤,٨٥	٤٦,٠	٥٣,٣
لوكسمبورج	٥٢٧,٢	٢٨,١٠	٣٠,٢	٣٦,٤
المملكة المتحدة	٦٥٦,٥	٦,٦٥	١٠,٨	١٤,٨٥

من الجدول السابق نستطيع استخلاص عدة نقاط هامة توضح العلاقة بين مؤشرات الاختلاف الثلاثة وبينها وبين الانحراف المعيارى وذلك على النحو التالى:

أولاً: أن ترتيب الدول الأربع حسب درجة الاختلاف لايختلف من مؤشر لآخر، فالترتيب فى كل حالة، باستثناء الانحراف المعيارى، هو: فرنسا، بلجيكا، لوكسمبورج، والمملكة المتحدة.

ثانياً: أن قيمة درجة الاختلاف تختلف اختلافاً واضحاً بين طرق المؤشرات الثلاثة. ويعنى ذلك أنه يجب عند توضيح درجة الاختلاف بين مجموعات البيانات أن نذكر الطريقة التى تم بواسطتها حساب قيمة الاختلاف (التباين) بينها. ويمكن القول أن الطريقة التى تستخدم كل من الانحراف الربيعى والوسيط تعطى قيماً أقل من الطريقتين الأخريتين، وذلك عند إهمال قيمة مؤشر الاختلاف لدولة بلجيكا التى تؤثر على الترتيب السابق لدرجة الاختلاف. ويرجع السبب فى ذلك إلى أن عيوب كل من الوسيط والانحراف الربيعى

التى ذكرناها سابقاً يتضمنها أيضاً مؤشراً لاختلاف الذى يعتمد عليهما فى حسابه. ونستطيع من الجدول إيجاد علاقة اعتبارية تربط بين معامل الاختلاف ومؤشر الاختلاف النسبى، إذ يلاحظ أن معامل الاختلاف أكبر من مؤشر الاختلاف النسبى وأن الفرق بين قيم كل منهما قد يصل إلى ٢٥٪ وهى بذلك تشبه العلاقة التى أوضحناها سابقاً بين الانحراف المتوسط والانحراف المعيارى.

ثالثاً: أن ترتيب الدول الأربع على أساس الانحراف المعيارى فقط يختلف عن الترتيب السابق مع مؤشرات الاختلاف الثلاثة. فمثلاً على الرغم من أن الانحراف المعيارى للمملكة المتحدة يعد الثانى من حيث الترتيب بينما نجد أن قيمة معامل الاختلاف لها تعد الأقل بين قيم معامل الاختلاف للدول الثلاث الأخرى. وإن دل هذا على شئ فإنما يدل على أن إهمال إدخال المتوسط الحسابى فى الحساب يؤثر تأثيراً كبيراً فى تقييم النتائج، أو بمعنى آخر إذا أردنا تقييم الاختلاف بالنسبة للمتوسط الحسابى فإن المقارنة تكون أكثر وضوحاً من مقارنة الانحرافات المعيارية نفسها.

القيم (الوحدات) المعيارية Standard Values (Units)

سبق أن ذكرنا أن معامل الاختلاف يستخدم كمقياس إحصائى للمقارنة بين تشتتى مجموعتين من البيانات العينية. أما إذا أردنا المقارنة بين قيمتين فى مجموعتين من البيانات العينية. أما إذا أردنا المقارنة بين قيمتين فى مجموعتين مختلفتين فإننا يجب أن نقارن بين موضع كل من هاتين القيمتين فى التوزيع الخاص بهما عن طريق إيجاد بعدهما عن المتوسط الحسابى بدلالة وحدات من الانحراف المعيارى. ويتم ذلك بقسمة انحراف كل قيمة عن متوسطها الحسابى على انحرافها المعيارى، ويمكن أن يكتب ذلك بالصيغة التالية:

$$\text{القيمة المعيارية} = \frac{س - \bar{س}}{ع}$$

حيث س هي قيمة المتغير، س هي متوسط قيم المتغير، عـ هي الانحراف المعيارى لقيم المتغير.

وتستخدم القيم المعيارية أيضاً في المقارنة بين المتغيرات التي تقاس على مقايس مزدوجة، كأن تكون قياسات المسافات بالكيلو مترات والأميال، أو تكون قياسات درجة الحرارة بالدرجات المئوية والفهرنهيتية، أو تكون قياسات المطر بالسنتيمترات والبوصات، وذلك للوقوف على معرفة توزيع وتشتت قيم المتغير في كل من حالات القياس المختلفة. فمثلاً إذا كانت لدينا المتوسطات الشهرية لدرجة الحرارة في المراصد الجوية في اقليم ما خلال شهور فصل الصيف (يونيو - أغسطس) وأردنا معرفة أى الشهور أنسب في درجة حرارته في هذا المرصد بالنسبة لبقية مواقع المراصد الأخرى، فإنه يمكن حساب القيمة المعيارية لكل شهر من الشهور ونقارنها مع بعضها البعض وذلك على النحو التالي:

	يونيو	يوليو	اغسطس
المعدل الشهرى في مراصد الاقليم	٢٥,٤	٢٦,٧	٢٦,٦
درجة الحرارة الشهرية في المرصد	٢٤,١	٢٥,٧	٢٦,٤
الانحراف المعيارى	١	١,٥	٢

وتكون القيمة المعيارية لكل شهر على حدة هي:

القيمة المعيارية لشهر يونيو = $\frac{٢٤,١ - ٢٥,٤}{١} = \frac{-١,٣}{١} = -١,٣$

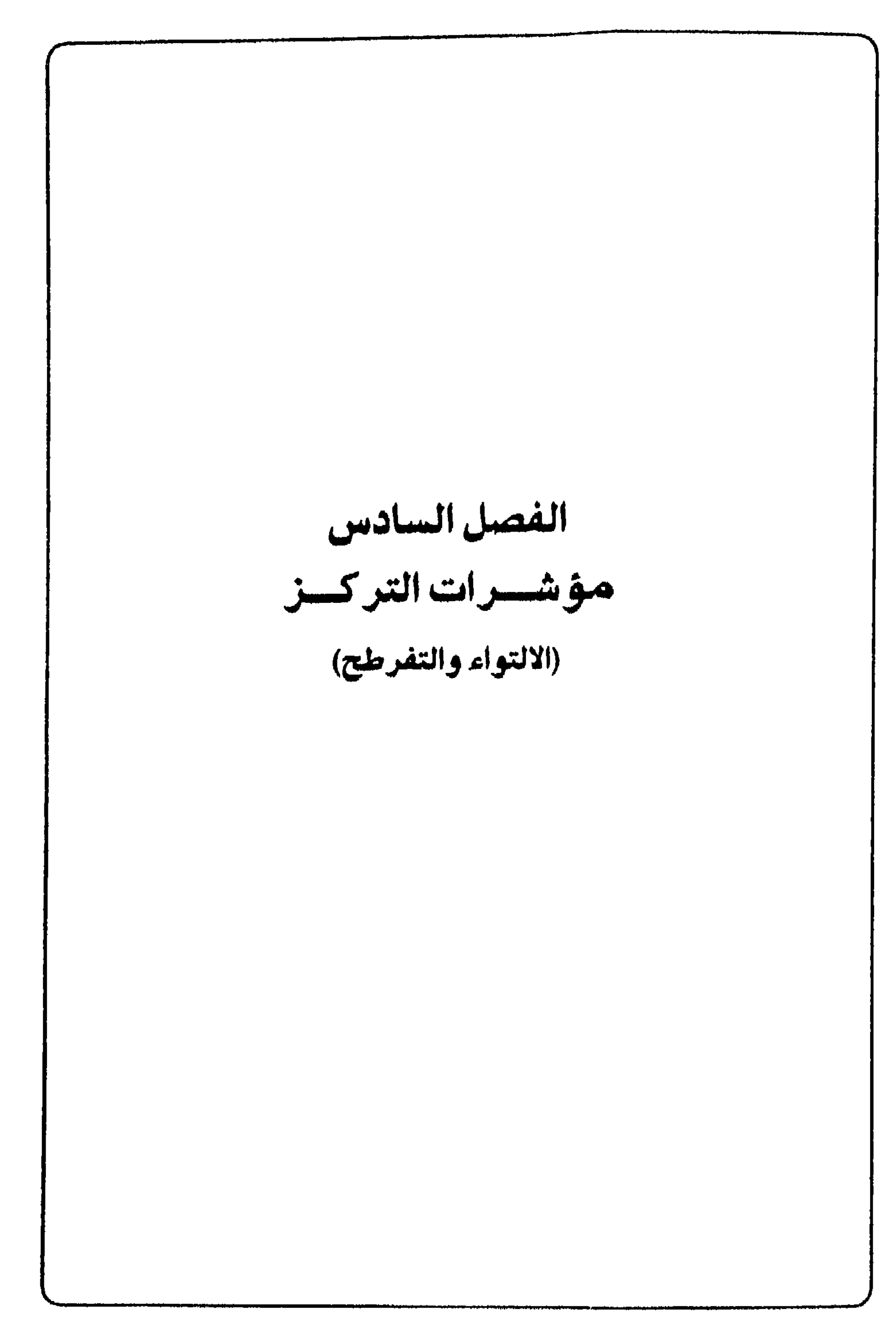
القيمة المعيارية لشهر يوليو = $\frac{٢٥,٧ - ٢٦,٧}{١} = \frac{-١}{١,٥} = -٠,٦٦$

القيمة المعيارية لشهر أغسطس = $\frac{٢٦,٤ - ٢٦,٦}{٢} = \frac{-,٢}{٢} = -٠,١$

وبما أن القيمة المعيارية لشهر أغسطس (-١,٠) هى أعلى القيم فإن درجة حرارة المرصد قيد البحث فى هذا الشهر كانت أفضل أو أكثر تحملاً من حرارة الشهرين الآخرين بالنسبة لبقية المراصد فى الاقاليم، ثم يأتى بعد ذلك شهر يوليو فشهر يونيو من حيث هذه الأفضلية.

والخلاصة أن مقاييس الانحراف ومؤشرات الاختلاف السابقة تعد ذات فائدة كبيرة فى التحليلات الاحصائية الجغرافية - أى فى المقارنات المختلفة بين المناطق والأقاليم - وذلك عن طريق تمثيل هذه المؤشرات على خرائط أو فى شكل رسوم بيانية. ومما تجدر الإشارة إليه هنا أنه لايوصى باستخدام كل من الانحراف الربيعى والانحراف المتوسط وما يرتبط بهما من مؤشرات للاختلاف (لتباين) بدرجة كبيرة فى التحليل الكمى للبحوث الجغرافية إذ يعانى كل منهما من نفس عيوب مقاييس المتوسطات التى تدخل فى تحديدها. ولكن يمكن القول بأن معظم الطرق والأساليب الكمية التى سيرد ذكرها فى متن الفصول القادمة تعتمد اعتماداً يكاد يكون كلياً على مقياس المتوسط الحسابى ومقاييس التباين (الانحراف المعيارى، ومعامل الاختلاف) والقيم المعيارية وهى المقاييس التى تستخدم، كما ذكرنا، كمؤشرات إحصائية لبيان طبيعة البيانات وكيفية توزيع مفرداتها، والتى سيبرز جوهرها ومضمونها بصورة واضحة عند استخدامها، أيضاً، فى التحليل الكمى للأمثلة الجغرافية فيما بعد.





الفصل السادس
مؤشـــرات التركـــز
(الالتواء والتفرطح)



مؤشـــرات التركـــز
(الالتواء والتفرطح)

سبق أن أوضحنا فى الفصلين السابقين كيفية حساب مقاييس النزعة المركزية (أو المتوسطات) وكذلك حساب مقاييس التشتت وفائدتها في وصف التوزيعات المختلفة ولإعطاء فكرة عن حجم البيانات الاحصائية وتشتتها حول متوسطها الحسابى. ولكن هذه المقاييس لا تكفى فى وصف التوزيعات ومقارنتها بعضها البعض، إذ أنه قد يتساوى توزيعان تكراريان من حيث المتوسط الحسابى والانحراف المعيارى ولكنهما يختلفان من حيث بعد المنحنى التكرارى للتوزيع عن التماثل. أو بمعنى آخر أن هذه المقاييس لا ينتج عنها معلومات عن خصائص ومميزات شكل Shap التوزيع التكرارى، بل تقيس درجة إتساع «عرض Width» المنحنى التكرارى للتوزيع. ويمكن تحديد بعد أو اقتراب التوزيع من التماثل من خلال شكل المنحنى التكرارى للتوزيع ومقارنته بمنحنى متماثل. كما يمكن كذلك تحديد شكل المنحنيات الوحيدة القمة من حيث تفرطحها أو درجة تدببها. فقد تتساوى بعض المنحنيات المتشابهة فى وجود قمة واحدة لها فى بعض الخصائص التى يمكن الحصول عليها بمقاييس النزعة المركزية والتشتت أو حتى بمقاييس عدم التماثل أو الإلتواء إلا أنها تختلف فى شكل قمتها. لكل ذلك فإننا سنلقى الضوء فى هذا الفصل على مؤشرين من المؤشرات الإحصائية التى تقيس اتجاهات تركز القيم هى: الإلتواء والتفرطح لما لهما من أهمية لا تقل عن أهمية تحديد المتوسط.

تحديد المتوسط الحسابى والتشتت فى تشخيص التوزيعات وتحديد خصائصها وملامحها.

الإلتواء Skewness

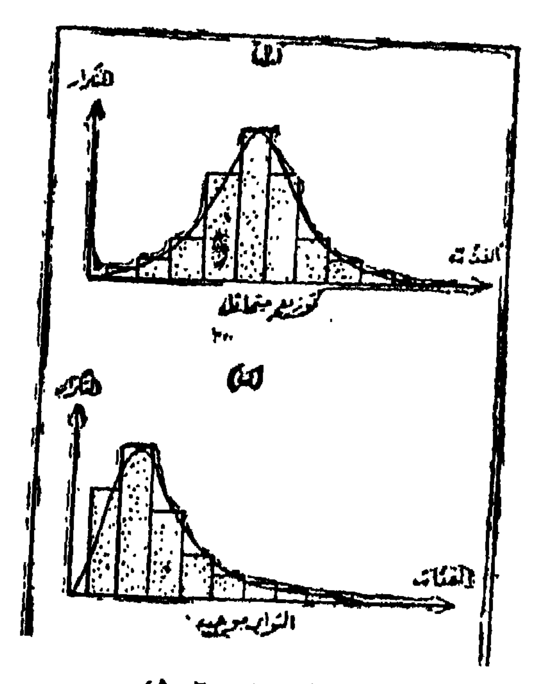
يعرف الإلتواء بأنه عبارة عن بعد لامدرج التكرارى أو المنحنى التكرارى عن التماثل حول المتوسط الحسابى للتوزيع، وهو بذلك يقيس اتجاه تركز القيم كما يحدد مناطق وجود بعض القيم المتطرفة فى التوزيع التكرارى. وباختصار فإنه يعطينا أسلوباً لتوزيع القيم وتحديد الامتداد الذى يتركز فيه القدر الأكبر منها على أحد جانبى المتوسط الحسابى ومقارنته بتوزيع القيم فى التوزيعات التكرارية المتماثلة. وتجدر الإشارة إلى أنه توجد بعض الظواهر الجغرافية التى تتوزع بياناتها توزيعاً متماثلاً أو قريباً جداً من التماثل حيث تتركز معظم القيم عند منتصف التوزيع، ولكن لا ينطبق ذلك على الكثير من ظواهر الجغرافية الطبيعية والتى تشبر طبيعياً بأثر فعل عامل طبيعى ينتج عنه بيانات تتوزع أو تتركز فى أحد أطراف التوزيع عنه فى الطرف الآخر مما يبعد ها عن التماثل، وكما سبق ذكره أنه عندما لخطنا البيانات الاحصائية على هيئة جداول توزيعات تكرارية ورسمنالها المدرجات التكرارية (الهيستوجرام)، فإن الأخيرة قد تأخذ الأشكال الثلاثة الآتية:

١- قد تتزايد التكرارات تدريجياً إلى أعلى حتى تصل إلى أكبر قيمة لها ثم تتناقص التكرارات إلى أسفل وبنسب تكاد تكون متساوية وتظهر هذه الخاصية بوضوح إذا أكثرنا من عدد الفئات وعدد المفردات وصغرنا من طول الفئة، وينعكس ذلك على شكل التوزيع الناتج حيث نحصل على مدرج تكرارى يسمى «مدرج تكرارى متماثل» (شكل رقم ٦-١أ) يوحى بأن المنحنى التكرارى المماثل له منحنى متماثل، إذ أن قمة المنحنى فى منتصف التوزيع تماماً، وأن طرفاه ينطبقان إلى درجة كبيرة على بعضهما عند المحور. ويعنى هذا أنه لا توجد قيم متطرفة أو شاذة سواء كانت كبيرة أو صغيرة تسبب ابتعاد أحد الأطراف عن الآخر، أو تؤدى إلى التواء التوزيع وعدم تحقيق تطابق طرفيه.

٢- قد يبدأ المدرج التكرارى بتركد كبير للتكرارات فى إطار الفئات الأولى

(الصغيرة) للتوزيع ثم يقل تركزها ويتناقص تدريجياً فى إطار الفئات الأخيرة (الكبيرة) وبصورة متطرفة مما يسبب عدم تطابق طرفى التوزيع، ويصبح شكل المدرج كما فى الشكل رقم (٦-١ب). ويقال إحصائياً أن المدرج التكرارى للتوزيع ملتو إلتواء موجباً أو ملتو ناحية اليمين حيث أن ذيل المنحنى الذى يمثل هذا التوزيع التكرارى يتجه ناحية اليمين من تأثير تطرف القيم فى الفئات الكبيرة.

٣- قد تبدأ التكرارا فى الفئات الأولى من التوزع صغيرة، مثل هذه التكرارات يمكن اعتبارها أيضاً متطرفة أو شاذة، ثم تزداد هذه التكرارات فجأة فى إطار الفئات الكبيرة. فى مثل هذه الحالة يقال أن التوزيع ملتو ناحية اليسار، وأن المنحنى الذى يمثل هذا التوزيع منحنى ملتو سالب حيث أن ذيل المنحنى أو الطرف الأيسر للتوزيع يتجه نحو اليسار أبعد من إتجاه الطرف الأيمن (شكل رقم ٦-١ حـ).



شكل رقم (٦-١)

أشكال المدرجات التكرارية وأنواع التوزيعات التكرارية التى تمثلها

والالتواء بهذا المفهوم السابق يعبر عن عدم التماثل فى التوزيعات التكرارية، ولذا فان تحديد مقدار ونوع الإلتواء يعد غاية فى الأهمية خصوصاً إذا علمنا أنه قد يكون هناك بعض التوزيعات التى تتساوى فى متوسطاتها الحسابية وأيضاً فى تشتتها فى الوقت الذى تختلف فيه تماماً فى التوائها. ومن هنا تميز التوزيعات عن بعضها، فقد يكون التواء التوزيعات فى إتجاه واحد ولكنه يختلف فى مقداره، أو قد تكون درجة الإلتواء فى التوزيعات متساوية ولكنها تختلف فى مقدار هذا الإلتواء، أو قد تكون درجة التوائها متساوية ولكنها تختلف فى النوع، ويمكن تحديد درجة الإلتواء (بسيط – متوسط – حاد) وأيضاً نوع الإلتواء (موجب – سالب) من خلال شكل المدرج أو المنحنى التكرارى للتوزيع ومقارنته بمدرج أو بمنحنى متماثل. إلا أن هذا الأسلوب لا يعطى قياساً دقيقاً لتحقيق هذا الغرض، ولذا فمن المستحسن إستخدام بعض المقاييس الكمية التى تقيس الإلتواء.

مقاييس الإلتواء:

لما كان التوزيع التكرارى المتماثل يتميز بانطباق كل من الوسط الحسابى والوسط والمنوال بعضها على بعض، فان وجود فرق بين هذه المقاييس إنما ينتج عنه التواء فى المنحنى. وبناء على ذلك فانه يمكن إستخدام الفروق بين قيم هذه المتوسطات لقياس للإلتواء. ونظراً لأن الفروق بين المتوسطات الثلاثة تتخذ شكل الوحدات المعيارية (القياسية) التى تختلف من توزيع إلى آخر فان هذه الفروق لا تقيس الإلتواء تماماً لأنه قد تكون الفروق كبيرة، والإلتواء بسيطاً، لزيادة تشتت البيانات، وقد تكون الفروق صغيرة والإلتواء حاداً لصغر تشتت لابيانات. ولذلك فاننا يجب أن ننسب الفروق بين المتوسط الحسابى والوسيط أو المتوسط الحسابى والمنوال إلى أحد مقاييس التشتت أو من نفس نوع مقاييس النزعة المركزية ونطلق على المقياس الناتج إسم «معامل الإلتواء». وتجدر الإشارة إلى أن معامل الإلتواء يجب أن يحقق شرطين أساسيين هما:

١- أن يكون هذا المعامل مساوياً للصفر وذلك بالنسبة للتوزيعات المتماثلة.

٢- أن لا تكون قيمة المعامل ذات تمييز معين، أو لا تتوقف على الوحدات التى

تقاس بها قيم المتغير. أو بعبارة أخرى أن تكون قيمته عدداً بحتاً. وعليه يمكن استعراض المقاييس الشائعة لحساب معامل الإلتواء فيمايلي:

$$\text{(١) معامل الإلتواء} = \frac{\text{الوسط الحسابي} - \text{المنوال}}{\text{الإنحراف المعياري}} \quad \text{أو}$$

$$\text{م.ت}_{١} = \frac{\bar{\text{س}} - \text{و}}{\text{ع}}$$

وفي التوزيعات الملتوية ناحية اليمين يقع المتوسط على نفس جانب المنوال أو في إتجاه القيم الكبيرة، أي تكون قيمة المتوسط الحسابي أكبر من قيمة المنوال ويكون المعامل حينئذ موجباً، والعكس مع التوزيعات الملتوية ناحية اليسار يكون معامل الإلتواء سالباً (راجع شكل رقم: ٤-٥أ، ب، جـ).

كما أنه يمكن إستخدام المقياس التالي وهو المسمى بمعامل بيرسون Pearson للإلتواء:

$$\text{(٢) معامل بيرسون للإلتواء} = \frac{\text{٣ (المتوسط الحسابي} - \text{الوسيط)}}{\text{الإنحراف المعياري}}$$

$$\text{م.ت}_{١} = \frac{\text{٣ (}\bar{\text{س}} - \text{ط)}}{\text{ع}}$$

وما ذكر عن المعامل م.ت$_{١}$ يقال أيضاً عن المعامل م.ت$_{٢}$ إلا أن قيمة م.ت$_{٢}$ تنحصر فيما بين + ١، -١، كما أن هذه الصيغة أكثر دقة من صيغة م.ت$_{١}$ حيث أن المنوال أقل دقة من الوسيط في وصف البيانات لأنه لا يأخذ في إعتباره إلا القيم الأكثر تكراراً ويهمل باقي القيم.

ويلاحظ على المقياسين السابقين في حساب الإلتواء أنهما يعتمدان على المتوسط الحسابي والإنحراف المعياري. وكما ذكرنا آنفاً أنه قد يتعذر في بعض الأحيان حساب المتوسط وبصفه خاصة في التوزيعات المفتوحة، كما أن حساب الإنحراف المعياري يحتاج بدوره إلى عمليات حسابية طويلة. لكل ذلك استخدمت

فكرة الفرق بين الربيعين (الأعلى والأدنى) والوسيط فى تحديد مقدار الإلتواء (معامل الإلتواء) الذى يمكن حسابه اذن على النحو التالى:

$$\text{معامل الإلتواء} = \frac{(\text{الربيع الأعلى} - \text{الوسيط}) - (\text{الوسيط} - \text{الربيع الأدنى})}{(\text{الربيع الأعلى} - \text{الربيع الأدنى})}$$

أى أن:

$$\text{م.ت}_{\text{٣}} = \frac{(\text{ر}_{\text{٣}} - \text{ط}) - (\text{ط} - \text{ر}_{\text{١}})}{(\text{ر}_{\text{٣}} - \text{ر}_{\text{١}})}$$

ونظراً لأن الفرق بين قيمة الربيع الأعلى والوسيط يساوى الفرق بين الوسيط والربيع الأدنى فى التوزيعات المتماثلة. فإذا كان هناك فرق بين كل من المقدارين كان ذلك دليلاً على عدم تماثل التوزيع، ووجود بعض القيم الشاذة هى التى تسبب هذا الفرق، وبالتالى التواء المنحنى ناحية اليمين أو اليسار أى التواء المنحنى التواء موجباً أو سالباً. ومن المعلوم أيضاً أنه فى حالة التوزيعات المتماثلة يقع الربيعان على بعدين متساويين من الوسيط، بينما فى التوزيعات الملتوية يختلف بعدا الربيع الأعلى والربيع الأدنى عن الوسيط، وبذلك يكون الفرق بين بعديهما عنه - أيضاً - مقياساً للإلتواء.

ويسمى معامل الإلتواء المحسوب بالصيغة السابقة بمقياس بولى Bowley - ويتميز بأنه المقياس الوحيد الذى يمكن استنتاجه من الرسم بدون الالتجاء إلى حساب أى قيمة وذلك باستخدام المنحنى التكرارى المتجمع الصاعد أو النازل. ويعاب عل يالمقياس السابق أنه لا يأخذ فى اعتباره قيم المفردات قبل قيمة الربيع الأدنى وكذلك قيم المفردات بعد قيمة الربيع الأعلى.

ونظراً لأن قيمة الإلتواء فى التوزيعات المتماثلة، كما سبق أن ذكرنا، تساوى صفراً فإن هذه القيمة تتخذ أساساً لتقدير نوع ودرجة حدة أو شدة الإلتواء. فكلما قربت قيمة أى معامل من معاملات الإلتواء الثلاثة السابقة من الصفر، كلما دل ذلك على وجود التواء ولكنه التواء بسيط، أما إذا بعدت قيمة معاملات الإلتواء عن الصفر وقربت من +١ أو -١ فإن ذلك يدل على كبر درجة حدة الإلتواء. فإذا

كانت إشارة معامل الإلتواء إشارة موجبة فإن ذلك يكون دليلاً على وجود التواء موجب يبتعد فيه الطرف الإيمن للمنحنى الممثل للتوزيع عن الطرف الأيسر مما يدل بالتالى على وجود بعض القيم المتطرفة (أو قيم كبيرة) والتى تقع فى إطار الفئات الأخيرة وتركز باقى القيم فى إطار الفئات الأولى للتوزيع. أما إذا كانت إشارة معامل الإلتواء إشارة سالبة فإن ذلك يكون دليلاً على وجود التواء سالب يبتعد فيه الطرف الأيسر لمنحنى التوزيع عن طرفع الأيمن مما نستنتج منه وقوع بعض القيم الشاذة (قيم صغيرة) فى إطار الفئات الأولى وتركز باقى القيم فى إطار الفئات الأخيرة ن التوزيع التكرارى للظاهرة قيد البحث.

مثال: من توزيع تكرارى لمساحة عدد ١٠٠ من المزارع حسبت المقاييس الإحصائية الآتية:

	فدان
المتوسط الحسابى للتوزيع	٢٧,٣٠
الوسيط	٢٧,٢٤
المنوال	٢٧,٤٢
الربيع الأدنى	٢٤,٠٠
الربيع الأعلى	٣٤,٧٠
الانحراف المعيارى	٩,٨٠

وكان المطلوب استنتاج معامل الإلتواء بالطرق السابقة، فإن ذلك ويكون على النحو التالى:

(١) م.ت = $\left[\frac{٢٧,٤٢ - ٢٧,٣}{٩,٨}\right]$ = – ٠و٠٠٦

(٢) م.ت = $\left[\frac{(٢٧,٢٥ - ٢٧,٣)٣}{٩,٨}\right]$ = ٠,٠١٥

(٢) م.ت = $\left[\frac{(٢٥,٠ - ٢٧,٢٥) - (٢٧,٢٥ - ٣٤,٧)}{(٢٥,٠ - ٣٤,٧)}\right]$ = ٠,٠١٤

وكما هو ملاحظ فإننا حصلنا على نتائج مختلفة لمعامل الالتواء من حيث نوع الإلتواء، إلا أنها تتفق جميعاً على وجود التواء ولكنه ضئيل جداً ويدل ذلك على أن التوزيع قريب جداً من التماثل. ويرجع ذلك إلى أن الأساس الذى حسب عليه معامل الإلتواء يختلف من طريقه لأخرى، كما أن كل طريقة تلائم بيانات خاصة ولا تلائم سواها. ويجب أن نوجه النظر هنا إلى أنه عند مقارنة التواء توزيعات مختلفة - لابد من استخدام صيغة واحدة لإيجاد معامل الإلتواء حتى يكون أساس المقارنة موحداً.

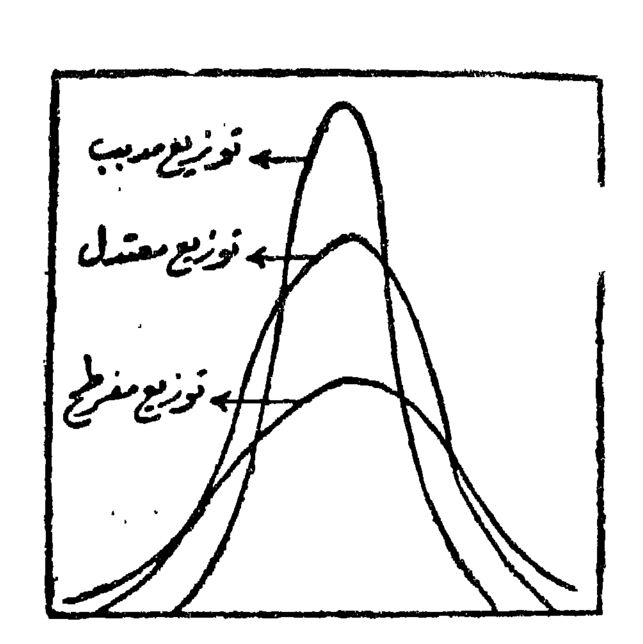
ويعتبر الإلتواء مقياساً إحصائياً له أهمية خاصة فى مجال الدراسات الجغرافية الكمية لأن معظم المتغيرات الجغرافية التى يمكن قياسها وجمع البيانات عنها تتصف توزيعاتها بأنها توزيعات شديدة الإلتواء مما يقف عائقاً أمام تطبيق الاختبارات الكمية البارامتريه Parameteric والتى تتطلب أن يكون التوزيع التكرارى لبيانات المتغير قيد البحث توزيعاً متماثلاً. هذا من ناحية ومن ناحية أخرى فإننا إذا استخدمنا أحد مقاييس لوصف الأخرى كالمتوسط الحسابى - بصفة خاصة - لوصف توزيع المتغير موضع الدراسة فأنه يكون مضللاً إذا كان بمفردة دون أن يقترن بتوضيح درجة ونوع التواء التوزيع. ولسوق مثالاً لتوضيح ذلك: إذا كان متوسط ما تنفقه الوحدة المحلية الواحدة بمحافظة ما على تنظيم النسل فى سنة ١٩٧١ هو ٧,٤٧ جنيهاً لكل ١٠٠٠ من السكان، فإذا اتخذت قيمة هذا المتوسط لذاتها فإنه يمكن افتراض أن نصف عدد الوحدات المحليه - تقريباً - بهذه المحافظة تنفق أكثر من هذه القيمة والنصف الآخر من الوحدات ينفق أقل من ٧٠٤٧ جنيهاً على تنظيم النسل بين سكانها. ولكن إذا تبين أن ربع عدد الوحدات هى التى تنفق على هذا الغرض أكثر من المتوسط ٧,٤٧ جنيهاً فأننا نتوقع أن العدد الباقى من الوحدات، وهو أكثر من النصف، ينفق أقل من المتوسط أولا ينفق شيئاً على تنظيم النسل بين السكان فى عام ١٩٧١ ..وفى مثل هذه الحالة - فان المتوسط لا يفيدنا كثيراً كمقياس إحصائى يعتمد عليه فى استخلاص المعلومات والنتائج. ولكن إذا قمنا

برسم المدرج التكرارى لمثل هذا التوزيع فان ذلك سوف يلقى ضوءاً سريعاً على حقيقة أن هذا التوزيع غير متماثل أو أنه ملتو إلتواء شديداً.

التفرطح Kurtosis

لا يقف تحليل المنحنيات البيانية التى تصف الكثير من المظاهر الخاصة ببيانات التوزيعات التكرارية على تحديد أو حساب كل من مقاييس النزعة المركزية أو التشتت أو حتى الإلتواء، بل يمتد إلى تحديد تفرطح أو درجة تديب المنحنيات البيانية الوحيدة القمة. ويعرف التفرطح إحصائياً بأنه ذلك المقياس الذى يقيس الامتداد الذى تتركز فيه القيم Values فى أحد أجزاء التوزيع التكرارى. فمثلاً إذا كانت إحدى فئات التوزيع أو مجموعة من الفئات المتجاورة تحتوى على نسبة كبيرة من تكرارات القيم داخل التوزيع، فإن هذا يعنى أن التوزيع مفرطح بدرجة كبيرة. ولكن درجة تفرطح أو شكل قمة المنحنيات البيانية تختلف من توزيع لآخر. فقد نجد قمة المنحنى البيانى لأحد التوزيعات عريضة أى مفرطحة، وهذا يعكس تركز القيم فى هذا التوزيع حول متوسطها الحسابى فى مدى كبير. ويسمى التوزيع فى هذه الحالة بتوزيع مفرطح Flat or Platykurtic. وقد نجد أن قمة التوزيع تبدو على شكل أكثر تدبباً، وهذا يعكس صغر مدى تركز القيم حول المتوسط الحسابى فيظهر شكل المنحنى البيانى ضيق فى الجزء العلوى ومتسع فى الجزء الأوسط. ويسمى التوزيع فى هذه الحالة توزيع مدبب Peaky or Leptokurtic. وقد نجد قمة منحنى التوزيع لاتبدو على شكل مفرطح أو مدبب وهذا يعكس تركز القيم حول متوسطها الحسابى بدرجة متماثلة. ويسمى التوزيع فى هذه الحالة توزيع متوسط التفرطح (توزيع متماثل) Mesokurtic (شكل رقم ٦-٢).

ويمكن التعرف على تفرطح أو تدبب المنحنيات البيانية بسهولة من خلال الشكل العام لها، غير أن هناك مقياس إحصائى يحدد درجة التفرطح فى التوزيعات بطريقة دقيقة اعتماداً على حساب مجموع القوة الرابعة لانحراف القيم عن متوسطها الحسابى مقسوماً على حاصل ضرب عدد القيم فى القوة الرابعة للإنحراف المعيارى. وإذا وضعنا ذلك فى صيغة جبرية فإنها تكون على النحو التالى:



شكل رقم (٦-٢)
أنواع التفرطح لمنحنيات التوزيعات التكرارية

$$\text{مقياس التفرطح} = \frac{\text{مجـ}(س - \bar{س})^{٤}}{ن\ عـ^{٤}} \qquad \text{(٦-٤)}$$

حيث

مجـ = مجموع

س = القيم المفردة كل على حدة

$\bar{س}$ = المتوسط الحسابى

ن = عدد القيم المكونه لسلسلة البيانات

عـ = الإنحراف المعيارى

وجدير بالذكر أن مقياس التفرطح غير شائع الإستخدام، أو أن إستخدامه ليس كما يجب أن يكون عليه، كمقياس إحصائى لوصف البيانات فى مجال الدراسات الجغرافية، على الرغم من الفوائد الهامة التى يمكن الحصول عليها ثم حسابه لمجموعة من البيانات، وهو بذلك يتشابه مع مقياس الإلتواء ومن المفيد أيضاً أن تشير إلى أن كثيراً من المتغيرات الجغرافة تتصف بشدة التواء وتدبب منحنياتها البيانية مما يجعل استخدام مقاييس الوصف الإحصائى الأخرى كالمتوسط الحسابى والإنحراف المعيارى أقل أهمية لما تعطية من إنطباعات مضللة عن خصائص توزيع بيانات تلك المتغيرات، ولو كانت هذه البيانات تخص عينات فإنها لا يمكن أن تكون ممثلة أو مسحوبة من مجتمعات إحصائية متماثلة التوزيع، كما إنها تكون غير ملائمة لتطبيق أساليب التحليل البارامترية التى تشترط أن تكون توزيعات البيانات متماثلة.

العزوم وقياس الإلتواء والتفرطح:

تستخدم العزوم Moments فى الإحصاء لبيان تماثل توزيع البيانات وكلمة «عزم» مشتقة من علم الاستاتسكا الذى يوضح أن قدرة القوة على تحريك جسم ما حول محور تتوقف على عاملين هامين هما: مقدار القوة، وبعد القوة عن المحور. وعلى ذلك فإن عزم القوة أو العزم حول المحور يعرف رياضياً على أنه حاصل ضرب مقدار القوة فى طول ذراعها (الذراع هو بعد خط عمل القوة عن مركز العزم أو البعد العمودى بين القوة وبين المحور). وتكون عزوم مجموعة من القوى مساوية مجموع حاصل ضرب كل قوة فى ذراع عزمها. وقياس العزم فى الإحصاء يختلف على قياسه فى الاستاتيكا (علم الثوابت) حيث يمكن اعتبار التكرارات فى التوزيعات مماثلة للقوة والقيمة المناظرة لهذه التكرارات (الفئة فى التوزيع) مطابقة لذراع العزم.

والمفهوم الإحصائى للعزم يتطلب تحديد النقطة التى نحسب عندها العزم. فقد يحسب العزم مثلا حول الصفر أو حول المتوسط الحسابى أو حول أى وسط فرضى آخر. إلا أن حساب العزوم حول المتوسط الحسابى (س) أصبح متعارفاً عليه كنقطة تحسب عندها العزوم.

وسنعرض فيمايلى الصيغ الجبرية لحساب العزوم حول المتوسط الحسابى للبيانات المبوبة.

١- العزم الأول (م) حول المتوسط الحسابى $= \dfrac{\text{١}}{\text{مجـ ك}}$ مجـ (م − سَ) ك

ولكن مجموع انحرافات القيم عن متوسطها الحسابى (م − سَ) = صفراً

∴ العزم الأول (م) حول المتوسط الحسابى $= \dfrac{\text{١}}{\text{مجـ ك}}$ مجـ (م − سَ) يساوى صفراً

٢- العزم الثانى ($\text{م}_{\text{٢}}$) حول المتوسط الحسابى $= \dfrac{\text{١}}{\text{مجـ ك}}$ مجـ $(\text{م} - \text{سَ})^{\text{٢}}$ك،
$=$ $\text{ع}^{\text{٢}}$ = التباين.

٣- العزم الثالث ($\text{م}_{\text{٣}}$) حول المتوسط الحسابى $= \dfrac{\text{١}}{\text{مجـ ك}}$ مجـ $(\text{م} - \text{سَ})^{\text{٣}}$ك،

وحيث أن م = مركز الفئة، (م − سَ) = الإنحراف عن المتوسط الحسابى = ح فإن

$$\text{العزم الثالث } (\text{م}_{\text{٣}}) = \left(\frac{\text{مجـ ح}^{\text{٣}}\text{ك}}{\text{مجـ ك}}\right) - \text{٣}\left(\frac{\text{مجـ ح}^{\text{٢}}\text{ك}}{\text{مجـ ك}} \times \frac{\text{مجـ ح ك}}{\text{مجـ ك}}\right) + \text{٢}\left(\frac{\text{مجـ ح ك}}{\text{مجـ ك}}\right)^{\text{٣}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (\text{٦} - \text{٨})$$

٤- العزم الرابع ($\text{م}_{\text{٤}}$) حول المتوسط الحسابى $= \dfrac{\text{١}}{\text{مجـ ك}}$ مجـ $(\text{م} - \text{سَ})^{\text{٤}}$ك ………………… (٦ − ٩)

أو

$$(\text{م}_{\text{٣}}) = \left(\frac{\text{مجـ ح}^{\text{٤}}\text{ك}}{\text{مجـ ك}}\right) - \text{٤}\left(\frac{\text{مجـ ح}^{\text{٣}}\text{ك}}{\text{مجـ ك}} \times \frac{\text{مجـ ح ك}}{\text{مجـ ك}}\right) + \text{٦}\left(\frac{\text{مجـ ح}^{\text{٢}}\text{ك}}{\text{مجـ ك}}\right) \times \left(\frac{\text{مجـ ح ك}}{\text{مجـ ك}}\right)^{\text{٢}} - \text{٣}\left(\frac{\text{مجـ ح ك}}{\text{مجـ ك}}\right)^{\text{٤}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (\text{٦} - \text{١٠})$$

ويمكن الاستطراد بنفس الطريقة السابقة للحصول على العزوم الأعلى حول المتوسط الحسابى، ولكننا عادة لانحتاج فى الدراسات الجغرافية العملية، وبصفة خاصة عند تحليل الرواسب المفتتة، إلى عزم أعلى من العزم الرابع.

وتجدر الإشارة هنا إلى أن العزم الثالث فى صورته السابقة (معادلة ٦-٨) يمكن اعتباره مقياسا دقيقا للإلتواء وحيث أن قيمة هذا العزم تساوى صفر للتوزيعات المتماثلة. أو بمعنى آخر إذا كان العزم الثالث يساوى صفرا فإن الإنحرافات السالبة تكون مساوية للإنحرافات الموجبة. وفى هذه الحالة فإن قيمة الإلتواء تساوى صفراً. أما فى التوزيعات غير المتماثلة فقد تكون قيمة العزم الثالث سالبة وهذا يعني وجود التواء فى التوزيع ناحية اليسار، أما إذا كانت قيمته موجبة فإن هذا يعنى وجود إلتواء فى التوزيع ناحية اليمين. وكلما كانت قيمة العزم الثالث قريبة من الصفر كلما كان منحنى التوزيع قريبا من التماثل، أما إذا كانت قيمته كبيرة (موجبة أو سالبة) كان المنحنى ملتو بشدة.

ولما كان العزم الثالث مقيساً بمكعب الوحدات المعيارية (القياسية) الأصلية فلابد فى حساب الإلتواء عن نسبة هذا العزم إلى أحد مقاييس التشتت أو الإختلاف مثل التباين (مربع الإنحراف المعيارى) للتوزيع بعد تكعيب الأخير وعلى ذلك فإن:

$$\text{معامل الإلتواء} = \frac{(\text{العزم الثالث})^{2}}{(\text{التباين})^{3}} = \frac{(\text{العزم الثالث})^{2}}{(\text{العزم الثانى})^{3}}$$

$$م.ت_{٤} = \frac{(م_{٣})^{2}}{(م_{٢})^{3}} \quad \text{......................} \quad (١٦ - ١)$$

كما تستخدم فكرة لعزم الرابع لقياس درجة التفرطح فى التوزيعات بطريقة دقيقة وذلك بقسمة العزم الرابع حول المتوسط الحسابى للتوزيع على مربع الإنحراف المعيارى له، ويمكن تحويل ذلك إلى صيغة إحصائية على النحو التالى:

$$\text{درجة التفرطح} = \frac{\text{العزم الرابع}}{(\text{التباين})^{٢}} = \frac{\text{العزم الرابع}}{(\text{العزم الثانى})^{٢}} ،$$

$$= \frac{م_{٤}}{(م_{٢})^{٢}} \quad \text{.....................} \quad (٦ - ١٢)$$

ويعزى السبب فى استخدام العزم الرابع (الذى يحتاج إلى رفع انحرافات القيم عن متوسطها الحسابى إلى القوة الرابعة) إلى وجود بعض القيم المتطرفة التى تمثل التوزيع، أما إذا لم توجد هذه القيم فإن العزم الرابع يعطى نتائج منخفضة، وبالقيمة على مربع العزم الثانى فإن قيمة التفرطح ستكون منخفضة إيضاً كما يرجع السبب فى القسمة على مربع العزم الثانى إلى التخلص من وحدات القياس والحصول على مقياس نسبى، وبحساب قيمة التفرطح للتوزيعات المتماثلة وجد أنها تساوى (٣) واعتبر التوزيع المتماثل متوزيع متوسط التفرطح أما إذا كان مقدار التفرطح أقل من (٣) فإن منحنى التوزيع يعد متفرطحاً منخفض القمة، بينما إذا كانت قيمة التفرطح أكبر من (٣) فان ذلك يعكس وجود منحنى مدببا يرتفع عن مستوى المنحنى المتماثل (شكل رقم ٦-٢). وفى النوع الأخير من المنحنيات تتركز القيم بشدة حول المتوسط الحسابى للتوزيع.

وتطبيقاً لما سبق يمكن حساب كل من قيمتى الإلتواء والتفرطح عن طريق استخدام العزوم من الجدول التالى (جدول رقم ٦-١) الذى يبين التوزيع التكرارى لمائة من الحيازات الزراعية فى أحد المراكز الإدارية حسب المساحة لهذه الحيازات

جدول رقم (٦-١)

حساب الإلتواء والتفرطح لتوزيع الحيازات الزراعية فى أحد المراكز الإدارية حسب مساحة الحيازة بإستخدام العزوم

مساحات	التكرارات	مراكز	ح	(ل/ح)	(ل/ح) ك	(ل/ح)^٢ ك	(ل/ح)^٣ ك	(ل/ح) ك
٥ -	٣	٧,٥	٢٠-	٤-	١٢-	٤٨	١٩٢-	٧٦٨
١٠-	٩	١٢,٥	١٥-	٣-	٢٧-	٨١	٢٤٣-	٧٢٩
١٥-	١٣	١٧,٥	١٠-	٢-	٢٦-	٥٢	١٠٤-	٢٠٨
٢٠-	١٦	٢٢,٥	٥-	١-	١٦-	١٦	١٦-	١٦
٢٥-	٢٠	٢٧,٥	صفر	صفر	صفر	صفر	صفر	صفر
٣٠-	١٥	٣٢,٥	٥	١	١٥	١٥	١٥	١٥
٣٥-	١٣	٣٧,٥	١٠	٢	٢٦	٥٢	١٠٤	٢٠٨
٤٠-	٨	٤٢,٥	١٥	٣	٢٤	٧٢	٢١٦	٦٤٨
٤٥-	٣	٤٧,٥	٢٠	٤	١٢	٤٨	١٩٢	٧٦٨
المجموع	١٠٠				٨١- ٧٧+ ٤-	٣٨٤	٥٥٥+ ٥٢٧+ ٢٨-	٣٣٦٠

ح = الإنحراف عن المتوسط (م – سَ)

ل = طول الفئة

ويكون حساب كل من العزوم الثلاثة حول المتوسط الحسابى كمايلى:

١– حساب العزم الثانى حول المتوسط الحسابى:

$$\text{م}_{٢} = \frac{١}{\text{ك}}\ \text{مجـ}\ \left(\frac{\text{ح}}{\text{ل}}\right)^{٢}\ \text{ك}$$

$$= \frac{١}{١٠٠} \times ٣٨٤ = ٣{,}٨٤ \text{ من الوحدات المربعة}$$

ولما كانت الوحدة المستخدمة (طول الفئة) هى ٥ أفدنة

$$\text{فإن : م}_{٢} = ٣{,}٨٤ \times (٥)^{٢} = ٩٦$$

٢– حساب العزم الثالث حول المتوسط الحسابى :

$$\text{م}_{٣} = \frac{\text{مجـ}\ (\text{ح}/\text{ل})^{٣}\ \text{ك}}{\text{مجـ ك}} - ٣ \times \frac{\text{مجـ}\ (\text{ح}/\text{ل})^{٢}\ \text{ك}}{\text{مجـ ك}} \times \frac{\text{مجـ}\ (\text{ح}/\text{ل})\ \text{ك}}{\text{مجـ ك}} + ٢\left(\frac{\text{مجـ}\ (\text{ح}/\text{ل})\ \text{ك}}{\text{مجـ ك}}\right)^{٣}$$

$$= \frac{-٢٨}{١٠٠} - ٣ \times \frac{٨٤٣}{١٠٠} \times \frac{-٤}{١٠٠} + ٢\left(\frac{٤٠}{١٠٠}\right)^{٣}$$

$= ٠{,}١٨٠٧$ من الوحدات المكعبة

ولما كانت الوحدة المستخدمة (طول الفئة) هى ٥ أفدنة

$$\text{فإن م}_{٣} = ٠{,}١٨٠٧ \times (٥)^{٣} = ٢٢{,}٥٩$$

٣– ويكون العزم الرابع حول المتوسط الحسابى هو :

$$\text{م}_4 = \frac{\text{مجـ}(\text{ل}/\text{ح})^4\,\text{ك}}{\text{مجـ ك}} - 4 \times \frac{\text{مجـ}(\text{ل}/\text{ح})^3\,\text{ك}}{\text{مجـ ك}} \times \frac{\text{مجـ}(\text{ل}/\text{ح})\,\text{ك}}{\text{مجـ ك}}$$

$$+\ 6 \times \frac{\text{مجـ}(\text{ل}/\text{ح})^2\,\text{ك}}{\text{مجـ ك}} \times \left(\frac{\text{مجـ}(\text{ل}/\text{ح})\,\text{ك}}{\text{مجـ ك}}\right)^2 - 3\left(\frac{\text{مجـ}(\text{ل}/\text{ح})\,\text{ك}}{\text{مجـ ك}}\right)^4$$

$$= \frac{٣٣٦٠}{١٠٠} - ٤ \times \frac{-٢٨}{١٠٠} \times \frac{-٤}{١٠٠} + ٦\left(\frac{٣٨٤}{١٠٠}\right)\left(\frac{-٤}{١٠٠}\right)^2 - ٣\left(\frac{-٤}{١٠٠}\right)^4$$

= ٣٣,٥٩ من وحدات ذات القوة الرابعة (الأس الرابع)

$= ٣٣,٥٩ \times (٥)^4$

$= ٢٠٩٩٣,٧٥$

وبناء على النتائج السابقة يمكن حساب قيمة كل من إلتواء ودرجة التفرطح للتوزيع كمايلى:

(١) معامل الإلتواء = $\frac{(\text{العزم الثالث})^2}{(\text{العزم الثالث})^3} = \frac{(\text{م}_3)^2}{(\text{م}_2)^3}$

$$\text{م.ث}_4 = \frac{(٢٢,٥٩)^2}{(٩٦)^3}$$

$= ٠,٠٠٠٦$

وهذا يعنى أن إلتواء التوزيع موجب وضعيف جداً، أى أنه توزيع قريب جداً من التماثل أو الاعتدال كما هو راضح من البيانات فى الجدول.

٢- درجة التفرطح = $\frac{\text{العزم الرابع}}{(\text{العزم الثانى})^2} = \frac{\text{م}_4}{\text{م}_2^2}$

$$= \frac{٢٠٩٩٣,٧٥}{(٩٦)^2}$$

$= ٢,٢٧٩$

وهذا يعنى أن التوزيع مفرطح ولكنه بدرجة لا تختلف كثيراً من تفرطح التوزيع المتماثل إذا أن قيمة التفرطح التى حصلنا عليها قريبة من ٣ (درجة تفرطح التوزيع المتماثل).

الباب الثالث
التقدير الإحصائي وأساليب المقارنة
Statistical Estimation and Comparison Technques



الباب الثالث

التقدير الإحصائى وأساليب المقارنة

مقدمة:

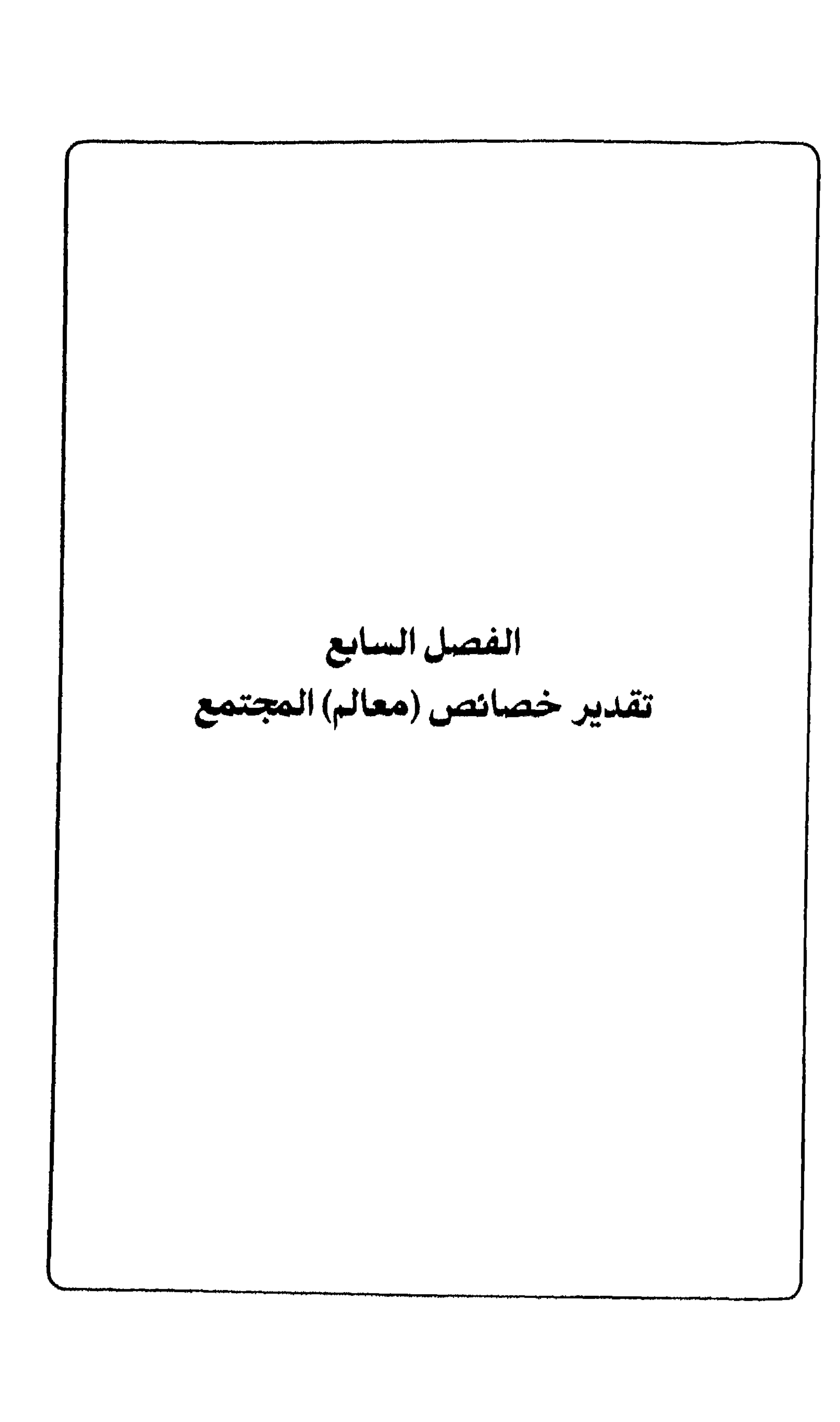
بعد دراستنا لأهم المقاييس الكمية المستخدمة فى وصف البيانات الجغرافية باستخدام أسلوب المعاينة، ودراسة بعض التوزيعات الاحتمالية لبعض المجتمعات الهامة لإلقاء الضوء على المفاهيم الأساسية والأساليب النظرية لتحديد احتمال صفات مفردات البيانات، وحيث أنه يتعذر فى كثير من الأحيان دراسة جميع مفردات المجتمع ولكن يمكن من الناحية العملية دراسة مفردات عينة مسحوبة من هذا المجتمع، فاهتمامنا الآن ينصب أولاً على مشكلة هامة فى الاستدلال الاحصائى Statistical Inference وهى مشكلة تقدير معالم المجتمع (مثل المتوسط الحسابى التباين ...) المجهولة من احصائيات العينة المعروفة.

وقد تتطلب المعالجة الاحصائية فى دراسة ظاهرة ما أن يضع الباحث بعض الفروض عن الخصائص (المعالم) المجهولة لمجتمع هذه الظاهرة، ويختار من بينها الفرض الذى لايكون معلوماً لديه، كما يحدد البديل المقابل للفرض قيد الاختبار. فمثلاً إذا أراد الباحث تقدير المتوسط العام لمجتمع ظاهرة ما عليه أولا أن يفترض قيمة معينة لهذه المعلمة المجهولة ثم يسحب عينة من بين مفردات المجتمع ليختبر ما إذا كانت الاحصائية المحسوبة من بيانات العينة تتفق أو لاتتفق مع فرضه السابق عن

معلمة المجتمع. ولاختبار مدى صحة الفرض أو قياس درجة الثقة فى التقدير تستخدم الاحتمالات ويسمى هذا الأسلوب الاحصائى باختبار الفروض Test of Hypotheses أو ما يعرف وأحياناً باختبار الدلالة أو المعنوية Test of Significance .

وبما أن جوهر البحث الجغرافي هو بيان، أو التعرف على الفروق والاختلافات المكانية Spatial differentiation فإنه يمكن الاستفادة والاستعانة بأساليب المقارنة الاحصائية الكمية للإجابة على كثير من التساؤلات التى تتصل بمدى دلالة أو أهمية هذه الفروق والاختلافات.

وفى هذا الباب سنناقش الأساليب الكمية البارامترية (المعلمية) والأساليب اللابارامترية (اللامعلمية) من خلال استخدامها للمقارنة بين خصائص مفردات مجموعة واحدة من البيانات وتوزيع نظرى لها، وللمقارنة بين مفردات مجموعتين أو بين ثلاث مجموعات أو أكثر من البيانات. وتجدر الإشارة هنا إلى أن صلاحية أو تطبيق أسلوب معين من هذه الأساليب الكمية يتوقف على نوع البيانات (نوعية، ترتيبية .... إلخ) المقيسة للظاهرة قيد البحث، وعلى خصائص الأسلوب الكمى نفسه.



الفصل السابع
تقدير خصائص (معالم) المجتمع



تقدير خصائص (معالم) المجتمع

رأينا فى الفصول السابقة أن دراسة المجتمعات تعتمد أساساً على الحصر الشامل لجميع مفردات المجتمع للتعرف على خصائص (معالم) هذا المجتمع. ويقصد بالمجتمع فى الدراسات الكمية، كما سبق أن ذكرنا، كل المفردات التى يتكون منها هذا المجتمع والتى تتصف بواحدة أو أكثر من الصفات المميزة المشتركة، وأن أية قيمة تحسب من توزيع المجتمع لدراسة خصائصه تسمى معلمة Paraneter فالمتوسط الحسابى، التباين، والانحراف المعيارى هى معالم لهذا المجتمع Population parameters.

وكما ذكرنا أن دراسة المجتمعات عن طريق أخذ كل مفردات المجتمع تعتبر من الأمور غير اليسيرة التى تحتاج إلى وقت طويل ومجهود كبير. هذا إلى جانب أنه فى كثير من الأحيان تتعذر دراسة كل المجتمع إذ يكون حجم المجتمع غير محدود Infinite أى لايمكن حصر جميع مفرداته، مثل مجتمع إنتاج سلعة من نوع معين، مما يفرض على الباحث القيام بفحص جزء من هذا المجتمع أو «عينة» منه. كذلك قد تتعذر دراسة كل المجتمع المحدود finite أى الذى يمكن حصر جميع مفرداته وذلك لأسباب اقتصادية أو عملية تقف أمام اتباع أسلوب الحصر الشامل لمعرفة القيمة الحقيقية لمعلمة المجتمع. فمثلاً إذا أردنا معرفة متوسط عمر المصباح من إنتاج أحد المصانع فى فترة معينة (المجتمع فى هذه الحالة هو مجتمع محدود ويتكون من الكمية المنتجة من المصابيح) فإنه يتعين علينا إضاءة كل مصباح من

إنتاج المصنع حتى يحترق لمعرفة عمره وبذلك نتمكن من معرفة متوسط عمر المصباح فى المصنع كله. أى أنه لمعرفة معلمة مجتمع المصابيح نضطر إلى اتلاف جميع مفرداته وهذا غير ممكن عملياً كما أنه يكون مكلف اقتصادياً. لذلك فإنه من الأوفق أن تأخذ عينة من هذه المصابيح الكهربائية وتترك مضاءة حتى تحترق ثم نستخدم قيمة متوسط عمر المصابيح فى العينة (احصائية العينة) كتقدير غير متحيز لمتوسط المجتمع (معلمة المجتمع). نخلص من ذلك أنه فى كثير من الأحيان لانستطيع معرفة القيمة الحقيقية بصفة مؤكدة لمعلمة المجتمع قيد البحث عن طريق الحصر الشامل ولذا فإننا نلجأ إلى تقديرها عن طريق اختيار عينة عشوائية من هذا المجتمع وحساب قيمة تقديرية لهذه المعلمة من بيانات العينة.

أنواع التقدير:

هناك نوعان من التقديرات لمعالم المجتمع هما: تقدير النقطة وتقدير الفترة الذى قد يكون مصحوباً بدرجة الثقة فى صحة التقدير. ويعبر تقدير النقطة Point Estimate عن معلمة المجتمع بقيمة واحدة كقيمة المتوسط الحسابى لعينة الذى يتخذ كتقدير غير متحيز أو تقدير قريب جداً من المتوسط الحقيقى للمجتمع. أما تقدير الفترة Interval Estimate أو فترة الثقة Confidence Interval فيعبر عن مدى معين من القيم بحيث يشتمل هذا المدى على قيمة المتوسط العام (أو المعلمة) للمجتمع. ولتوضيح ذلك نذكر أنه إذا قيست مسافة على خريطة فكانت ٥٢,٨ سنتيمتراً أو إذا قدر عمر قطعة أثرية بألف سنة فإننا فى هذه الحالة بصدد تقدير نقطة لطول المسافة ولعمر القطعة الأثرية. ومن الناحية الأخرى إذا ذكرنا أن المسافة هى ٥٢,٨ ± ٣, سنتيمتراً أى أن المسافة تقع بين ٥٢,٥ سنتيمتراً و ٥٣,١ سنتيمتراً باحتمال قدره ٠,٩٥، أو إذ قدر عمر قطعة أثرية بين ١٠٠٠ سنة و ١٥٠٠ سنة باحتمال قدره ٠,٩٥، فإننا نعطى تقديراً بفترة لطول المسافة أو لعمر القطعة الأثرية واحتمال أن يقع كل منهما فى أى نقطة منها. ونظراً لأن تقدير الفترة يعطى قيمة احتمال وقوع (٠,٩٥) وكذلك عدم وقوع (٠٥,) طول المسافة

أو العمر فى هذه الفترة، ولذلك فإننا نطلق على الفترة (٥٢,٥ ، ٥٣,١ سنتيمتراً أو ١٠٠٠ ، ١٥٠٠ سنة) إسم «فترة ثقة ٩٥٪» لطول المسافة ولعمر القطعة الأثرية. وبصفة عامة فإن التقديرات بفترة تشير إلى معنوية أو دقة التقدير، وبالتالى فإنها تفضل على التقدير بنقطة لمعلمة من معالم المجتمع.

تقدير فترة الثقة لمعالم المجتمع:

كثيراً ما يكون هناك مجتمع لاتعرف معالمه (المتوسط الحسابى «مَ» أو الانحراف المعيارى «ع») ونجد أنه بينما نريد معرفة بعض أو كل هذه المعالم فإننا لانستطيع تحديد هذه المعالم تحديد دقيقاً ومؤكداً وذلك لأسباب عملية أو اقتصادية. وفى هذه الحالة نلجأ إلى تقدير معالم المجتمع الأصلى من خلال إحصائيات عينة عشوائية تسحب من هذا المجتمع. على أنه يمكن القول أن هذا التقدير يتطلب معرفة طبيعة العلاقة بين التوزيع الأصلى للمجتمع بمعالمه المختلفة وتوزيع المعاينة للمتوسطات الحسابية بإحصائياته المختلفة.

ومن الناحية العملية لايمكننا سحب عدد كبير من العينات، لاعتبارات مالية وزمنية، ولكن يمكن سحب عينة واحدة ونحسب منها المتوسط الحسابى (سَ) ومنه يمكن تقدير لمتوسط الحسابى للمجتمع (مَ) بإنشاء فترة ثقة للمتوسط الحسابى المحسوب من العينة (سَ). فإذا كان حجم العينة كبيراً بدرجة كافية (ن أكبر من ٣٠) ومسحوبة عشوائياً من مجتمع – ليس بالضرورة أن يكون معتدلاً – له متوسط حسابى (مَ) وانحراف معيارى (ع) فإننا يمكن أن نتوقع، تبعاً لخصائص التوزيع المعتدل، أن نجد قيمة فعلية للمعلمة (مَ) تقع بالتقريب:

فى الفترة [سَ + ع ، سَ – ع] باحتمال ٦٨,٢٧ ٪

وفى الفترة [سَ + ٢ ع، سَ – ٢ ع] باحتمال ٩٥,٤٥ ٪

وفى الفترة [سَ + ٣ ع ، سَ – ٣ ع] باحتمال ٩٩,٧٣ ٪

أما إذا رجعنا إلى خصائص توزيع المعاينة للمتوسطات الحسابية المسحوبة

عشوائياً من المجتمع المراد تقدير المعلمة (مَ) له، فإننا يمكن أن نتوقع تبعاً لخصائص التوزيع العينى، الذى يتوزع بشكل تقريبى كالتوزيع المعتدل الذى متوسط حسابى مَ سَ وانحراف معيارى (خطأ معيارى) $\frac{ع}{\sqrt{ن}}$ ، أن نجد أن قيمة المعلمة (مَ) تقع بالتقريب:

في الفترة $\left[\bar{س} + \frac{ع}{\sqrt{ن}} \text{ ، } \bar{س} - \frac{ع}{\sqrt{ن}}\right]$ باحتمال ٦٨,٢٧ ٪

وفي الفترة $\left[\bar{س} + \frac{٢ع}{\sqrt{ن}} \text{ ، } \bar{س} - \frac{٢ع}{\sqrt{ن}}\right]$ باحتمال ٩٥,٤٥ ٪

وفي الفترة $\left[\bar{س} + \frac{٣ع}{\sqrt{ن}} \text{ ، } \bar{س} - \frac{٣ع}{\sqrt{ن}}\right]$ باحتمال ٩٩,٧٣ ٪

وتسمى الفترات ٦٨,٢٧ ٪، ٩٥,٤٥ ٪، ٩٩,٧٣ ٪ بفترات الثقة، أو مستوى الثقة Confidence level، (الاحتمال) لتقدير المعلمة (مَ)، وبالتالى فإن درجة عدم الصحة فى التقدير للفترة الثانية = (١٠٠ – ٩٥,٤٥ ٪) = ٤,٥٥ ٪ وعدم للصحة فى التقدير للفترة الثالثة = (١٠٠ – ٩٩,٧٣ ٪) = ٠,١٧ ٪ كما يطلق على حدود هذه الفترات (سَ ± ع، سَ ± ٢ ع، سَ ± ٣ ع) اسم حدود الثقة Confidence limits للمعلمة (مَ). أما حدود الثقة للفترات ٩٥ ٪، ٩٩ ٪ فهى على الترتيب: سَ ± ١,٩٦ ع، سَ ± ٢,٥٨ ع. وتسمى القيم ١,٩٦ ، ٢,٥٨ فى حدود الثقة بمعاملات الثقة Confidence Coefficients أو القيم الحرجة Critical Values ويرمز لها بالرمز (ز١). ويمكن الحصول على معاملات الثقة من مستوى الثقة المطلوب للتقدير، أو العكس، كما يلى:

مستوى الثقة (الاحتمال)	٦٨,٢٧	٪٨٠	٪٩٠	٪٩٥	٪٩٦	٪٩٨	٪٩٩
(ز)	١,٠٠	١,٢٨	١,٦٤٥	١,٩٦	٢,٠٥	٢,٣٣	٢,٥٨

من كل مما سبق يمكن القول أن إنشاء فترة الثقة يعتمد أساساً على الخصائص التى سبق ذكرها عن التوزيع الاحتمالى المعتدل والتوزيع الاحتمالى للمتوسطات الحسابية للعينات الذى سبق أن قلنا أن له توزيع معتدل إذا كان حجم العينة كبير بدرجة كافية بصرف النظر عن شكل التوزيع الأصلى للمجتمع المسحوب منه العينة، وأن المتوسط الحسابى للمتوسطات الحسابية لجميع العينات الممكنة يساوى تقريباً المتوسط العام (مَ)، كما أن الانحراف المعيارى للمتوسطات الحسابية يساوى الانحراف المعيارى للمجتمع مقسوماً على حجم العينة (مع إهمال معامل التصحيح أو «معامل بسل» Bessel's Correction فى العينات الكبيرة) أو ما يعرف بالخطأ المعيارى Standard Error وهو الخطأ الناجم عن احتمال بعد أو قرب خصائص العينة فى تمثيل معالم المجتمع، أى أنه عبارة عن مدى تفاوت متوسط العينة مثلاً عن متوسط المجتمع المسحوبة منه. وبما أن الانحراف المعيارى للمجتمع ثابت بينما حجم العينة متغير فإن ذلك يعنى أنه كلما زاد حجم العينة كلما صغرت قيمة الخطأ المعيارى، وكلما كان الخطأ المعيارى صغيراً كلما كانت قيمة متوسط العينة قريبة من قيمة متوسط المجتمع، وكان بالتالى تمثيل على معرفة الانحراف المعيارى للمجتمع فيمكن تقديره بسهولة. ولكن قد يحدث فى بعض الأحيان أن لايكون الانحراف المعيارى للمجتمع معلوماً. وللتغلب على هذه الصعوبة فإننا نلجأ إلى تقدير الخطأ المعيارى عن طريق استعاضة الانحراف المعيارى للمجتمع بالانحراف المعيارى للعينة.

وقد جرت العادة فى الأبحاث الجغرافية على إنشاء فترة بدرجة ثقة ٩٥٪ إلا أنه يمكن أيضاً إنشاء فترة بأية درجة ثقة. ونظراً لعدم التأكد من أن فترة الثقة المحسوبة تشتمل أو لاتشتمل على المتوسط الحقيقى (المعلمة) للمجتمع فإن تقديرنا يبنى على أساس عنصر الاحتمال إذ يمكن أن نتحكم فى نسبة الخطأ الذى قد تقع فيه باستخدام درجة ثقة معينة فى التقدير، فإذا استخدمنا درجة ثقة كبيرة فإن فترة الثقة المرتبطة بها تكون كبيرة. وبالطبع كلما زاد طول فترة الثقة كلما

قلت قيمتها العملية لذلك فإنه من المهم أن تكون فترة الثقة التى نقررها ذات فائدة عملية للبحث.

التقدير من احصائية (مقاييس) العينات:

ذكرنا من قبل أنه إذا كان لدينا مجتمعاً، ليس بالضرورة أن يكون توزيعه معتدلاً، متوسطة (مَ) وانحرافه المعيارى (ع) وسحبنا منه كل العينات الممكنة التى حجمها (ن) فإن توزيع المعاينة للمتوسطات الحسابية للعينات يقترب من التوزيع المعتدل الذى متوسطة (مَ) وانحرافه المعيارى (الخطأ المعيارى) $\frac{ع}{\sqrt{ن}}$. وكما نعلم من نظرية النهاية المركزية التى مؤداها أنه إذا كان لدينا عينة حجمها كبير بدرجة كافية (ن أكبر من ٣٠) ومسحوبة عشوائياً من مجتمع (ليس بالضرورة أن يكون معتدلاً تماماً) له متوسط حسابى (مَ) وانحراف معيارى $\frac{ع}{\sqrt{ن}}$ فإن قيمة المتوسط الحسابى للعينة (سَ) يمكن اعتبارها تقديراً غير متحيز لمتوسط (معلمة) المجتمع إذا ربطنا هذا التقدير بدرجة معينة للثقة أو بنسبة معينة للخطأ فى التقدير والتى على أساسها ننشئ فترة الثقة للمتوسط الحسابى المحسوب من العينة بدرجة ثقة ٩٥٪ أو ٩٩٪ (بمعنى أن احتمال أن تشتمل هذه الفترة المقدرة على المتوسط العام للمجتمع (مَ) يساوى ٩٥, أو ٩٩,) لابد أن تمتد هذه الفترة بين +١,٩٦ ، −١,٩٦ أو +٢,٥٨ ، −٢,٥٨ من الانحراف المعيارى للمتوسطات الحسابية أو من الخطأ المعيارى، ولذلك فإن:

حدود فترة الثقة للمتوسط الحسابى للمجتمع (مَ) =

$$سَ \pm ز_{\alpha^{1}/_{2}} \times \frac{ع}{\sqrt{ن}} \quad \ldots \quad \ldots \quad \ldots \quad \ldots (١٠ - ١)$$

فى حالة ما إذا كانت العينة مسحوبة من مجتمع غير محدود أو إذا كانت

المعاينة بارجاع من مجتمع محدود حيث $(\bar{س} + ز_{\alpha/2} \times \frac{ع}{\sqrt{ن}})$ هو الحد الأعلى لفترة الثقة، $(\bar{س} - ز_{\alpha/2} \times \frac{ع}{\sqrt{ن}})$ هو الحد الأدنى لفترة الثقة. أما إذا كانت المعاينة بدون ارجاع من مجتمع محدود حجمه ن فإن:

حدود فترة الثقة للمعلمة $(\bar{م})$ =

$$\bar{س} \pm ز_{\alpha/2} \frac{ع}{\sqrt{ن}} \times \sqrt{\frac{ن - ن}{ن - ١}} \quad \ldots \quad \ldots \quad \ldots \quad (١٠-٢)$$

وفي الحالات التي لايكون فيها الانحراف المعياري (ع) معلوماً يستعاض عنه بالانحراف المعياري للعينة (عـ) للحصول على حدود الثقة السابقة لتقدير متوسط المجتمع ويكون ذلك صحيحاً إذا كان حجم العينة أكبر من ٣٠ مفردة.

مثال (١)

إذا كان توزيع الأجور لعمال أحد المصانع يتوزع توزيعاً قريباً جداً من الاعتدال، وبأخذ عينة عشوائية حجمها ١٠٠ عامل من عمال هذا المصنع وجد أن متوسط الأجور في العينة هو ٧٠ جنيهاً في الشهر فأوجد فترة ثقة ٩٥ ٪ للمتوسط الحسابي لأجور العمال في هذا المصنع علماً بأن الانحراف المعياري لأجور العمال في المصنع هو ١٠ جنيهات.

بما أن الانحراف المعياري للمجتمع معلوماً فإن:

$$\text{الخطأ المعياري} = \frac{ع}{\sqrt{ن}} = \frac{١٠}{\sqrt{١٠٠}} = ١$$

ومن جداول التوزيع المعتدل المعياري نجد أن المساحة تحت المنحنى التي تشتمل على ٩٥ ٪ (درجة الثقة) من القيم تنحصر بين $ز_{\alpha/2}$ (أى ١٠ - ٩٥, = ٠٥, ÷ ٢ = ز ٠,٢٥) ± ١,٩٦.

∴ الحد الأدنى لفترة الثقة ٩٥٪ = $\bar{س}$ − ١,٩٦ × $\frac{ع}{\sqrt{ن}}$

= ٧٠ − ١,٩٦ × ١ = ٦٨,٤٠ جنيهاً

الحد الأدنى لفترة الثقة ٩٥٪ = $\bar{س}$ + ١,٩٦ × $\frac{ع}{\sqrt{ن}}$

= ٧٠ + ١,٩٦ × ١ = ٧١,٩٦ جنيهاً

وبذلك تكون (٦٨,٠٤)، (٧١,٩٦) هى فترة الثقة ٩٥٪ لمتوسط المجتمع (المعلمة مَ).

مثال ٢:

أرادت مصلحة الضرائب بمحافظة الإسكندرية معرفة متوسط الأرباح التجارية السنوية للمحلات الصغيرة لتقدير الضرائب المستحقة على أصحاب هذه المحلات فسحبت عينة عشوائية من ١٠٠ محل حسب منها المتوسط الحسابى للمبيعات السنوية فكان ١١٠٠ جنيه كما حسب الانحراف المعيارى لأرباح مجتمع المحلات التجارية فكانت ١٢٠ جنيه والمطلوب ايجاد فترة ثقة ٩٩٪ لمتوسط الأرباح لمجتمع هذه المحلات.

بما أن الانحراف المعيارى للمجتمع معلوماً فإن:

∴ الخطأ المعيارى = $\frac{ع}{\sqrt{ن}}$ = $\frac{١٢٠}{\sqrt{١٠٠}}$ = ١٢ جنيهاً

وحيث أن أكبر من ٣٠ فإننا نستخدم جداول المنحنى المعتدل المعيارى لإيجاد القيمة المقابلة لقيمة ز $_{a١/٢}$ فتكون :

ز $_{a١/٢}$ = ز ٠,٠٠٥ (أى : ١ − ٠,٩٩ = ٠,٠١ ÷ ٢) = ± ٢,٥٨

∴ الحد الأدنى لفترة ٩٩٪ = $\bar{س}$ − ٢,٥٨ × $\frac{ع}{\sqrt{ن}}$

= ١١٠٠ − ٢,٥٨ × ١٢ = ١٠٦٩,٠٤ جنيها

الحد الأعلى لفترة الثقة ٩٩٪ = $\bar{س}$ + ٢,٥٨ × $\frac{ع}{\sqrt{ن}}$

= ١١٠٠ + ٢,٥٨ × ١٢ = ١١٣٠,٩٦ جنيها

إذا نتوقع أن متوسط الأرباح للمحلات التجارية الصغيرة في محافظة الاسكندرية يقع في الفترة بين ١٠٦٩٠,٠٤ ، ١١٣٠,٩٦ جنيه بدرجة ثقة ٩٩٪.

مثال (٣):

لمعرفة متوسط استدارة الرواسب الحصوية على جزء من الشاطئ سحبت عينة عشوائية من هذه الرواسب حجمها ١٠٠ حصوة وحسب منها المتوسط الحسابي (٥٠ ملليمترا) والانحراف المعيارى (١٠ ملليمترا). والمطلوب حساب فترة الثقة ٩٥٪ لتقدير المتوسط العام لمجتمع الرواسب الحصوية على الشاطئ قيد الدراسة.

بما أن الانحراف المعيارى لمجتمع الرواسب الحصوية على الشاطئ غير معلوم، وأن حجم العينة أكبر من ٣٠ مفردة فإننا نحسب الخطأ المعيارى للتقدير باستخدام الانحراف المعيارى للعينة:

الخطأ المعيارى = $\frac{ع}{\sqrt{ن}}$ = $\frac{١٠}{\sqrt{١٠٠}}$ = ١

وباستخدام جداول المنحنى المعتدل المعيارى نجد أن:

$ز_{\alpha ١/٢} = (١ - ٠,٩٥ = ٠,٠٥ ÷ ٢ = ز_{a١/٢}) = \pm ١,٩٦$

∴ الحد الأدنى لفترة الثقة ٩٥٪ = ٥٠ - ١,٩٦ × ١ = ٤٨,٠٤ ملليمتراً

الحد الأعلى لفترة الثقة ٩٥٪ = ٥٠ + ١,٩٦ × ١ = ٥١,٩٦ ملليمتراً

وهذا يعنى أن متوسط استدارة الرواسب الحصوية على الشاطئ يقع بين الحدين (٤٨,٠٤، ٥١,٩٦ ملليمترا)، أو بمعنى آخر لا يقل عن ٤٨,٠٤ ملليمتراً ولايزيد عن ٥١,٩٦ ملليمتراً بدرجة ثقة ٩٥٪ وبنسبة خطأ مسموح به ٥٪.

وفى بعض الأحيان يرغمنا عنصر التكاليف فى جمع البيانات إلى سحب عينة صغيرة (ن أقل من ٣٠ مفردة) لتقدير معالم المجتمع الذى تمثله. وفى مثل هذه الحالات لا يمكن الاستعانة بنظرية النهاية المركزية إذ أن توزيع للقيم المعيارية للمجتمع غير معلوماً والعينة حجمها صغير فإنه يمكن استبدالة بالانحراف المعيارى للعينة بعد إجراء بعض التعديل على الأخير حتى نحصل على «أحسن تقدير للإنحراف المعيارى للمجتمع» لأن الخطأ المعيارى المحسوب من واقع الانحراف المعيارى للبيانات المشاهدة من العينة الصغيرة قد يختلف كثيراً عن الانحراف المعيارى للمجتمع الأصلى مما يؤثر على درجة الدقة فى التقدير والاستنتاج الاحصائى، وبذلك فإن:

أحسن تقدير للإنحراف المعيارى للمجتمع ($\hat{ع}$)= الانحراف المعيارى للعينة (د) × تصحيح «بل»، أى أن:

$$\hat{ع} = عـ \times \sqrt{\frac{ن}{ن - ١}}$$

ويكون الخطأ المعيارى للمتوسط الحسابى ($\bar{س}$) عبارة عن :

$$\text{الخطأ المعيارى} = \frac{\hat{ع}}{\sqrt{ن}}$$

أو = $\frac{\hat{ع}}{\sqrt{ن-١}}$

وتسمى القيمة (ن – ١) بعدد المتغيرات المستقلة الخطية التى يمكن تكوينها من ن من القيم المساهدة أو ما يعرف بدرجات الحرية.

وفى حالة العينات الصغيرة أيضاً تستخدم الإحصائية ت (t) بدلاً من (ز) لتقدير القيمة الإفتراضية لمعلمة المجتمع (مَ)، وذلك لأن الأولى تتميز بأن توزيعها ينتشر على مدى أوسع من التوزيع المعتدل، ومن ذلك نتوقع أن نحتاج إلى أكثر من خطأين معيارين لتحديد فترة الثقة ٩٥٪ لمتوسط المجتمع (١,٩٦ فى حالة التوزيع المعتدل المعيارى «ز»)، وتعتمد قيمة (ت) على حجم العينة أو بالأخرى على درجات الحرية (ن – ١) ولذلك فإن المتغير المعيارى (ت) ليس له توزيع إحتمال واحد كالمتغير المعيارى (ز)، ولكن له توزيع إحتمالى لكل قيمة من قيم درجات الحرية (من ١ إلى ٣٠). والجدول المختصر فى ملاحق الكتاب يبين قيمة (ت) عند مستويات معنوية مختلفة لمساحة الطرف الموجب للمنحنى وهو كاف لإستخدامه فى إنشاء فترات الثقة للعينات الصغيرة. فإذا كان حجم العينة = ١٠، ومستوى المعنوية (α) = ٠,٠٥ فإن قيمة (ت) التى تجعل مساحة كل طرف من طرفى المنحنى = ٢,٥٪ من المساحة الكلية تقع فى الصف تحت درجة الحرية ٩ (أى ١٠–١) وفى العمود ٠,٠٢٥ تساوى ٢,٢٦٢ أى أن ٩٥٪ من المساحة الكلية للمنحنى تنحصر بين (ت) + ٢,٢٦٢ (لاحظ أن القيمة المناظرة فى التوزيع المعتدل المعيارى (ز) = ± ١,٩٦). إلا أنه عندما يكبر حجم العينة تصبح قيمة $\hat{ع}$ قريبة جداً من ع، كما يقترب التوزيع الاحتمالى للمتغير (ت) ثم التوزيع الاحتمالى للمتغير الاحتمالى للمتغير المعتدل المعيارى (ز)،وهكذا يمكن استخدام التوزيع المعتدل كتقريب لتوزيع (ت) إذا كانت درجات الحرية تساوى أو أكثر من ٣٠. (ويلاحظ من جدول توزيع «ت» أن القيمة الأخيرة عند «ن –١» = α هى نفسها قيمة «ز» فى المنحنى المعتدل.) وبذلك يقتصر استخدام توزيع (ت) على

الحالات التى تكون فيها حجم العينة أقل من ٣٠ والانحراف المعيارى للمجتمع غير معلوم.

ويمكن إنشاء فترة ثقة لتقدير متوسط المجتمع (مَ) بنفس الطريقة المستخدمة سابقاً فى حالة العينات الكبيرة، إلا أننا نستخدم فى هذه الحالة قيم (ت) بدرجات حرية (ن − ١) بدلاً من قيمة (ز). وتكتب صيغة حدود فترة الثقة لمعلمة المجتمع (مَ) كالتالى:

$$\text{حدود فترة الثقة} = \bar{\text{س}} \pm \text{ت} \times \frac{\hat{\text{ع}}}{\sqrt{\text{ن} - ١}} \quad \ldots \quad \ldots \quad \ldots \quad (١٠ - ٣)$$

مثال (٤):

أخذت عينة من مجموعة من الروافد ذات الرتبة الأولى فى أحد الأحواض النهرية مكونة من ٢٥ رافداً لدراسة انحدار جوانبها فوجد أن متوسط الانحدارات هو ٢٠ درجة بانحراف معيارى ٥ درجات، والمطلوب تقدير متوسط انحدار جوانب كل الروافد من نفس الرتبة وذلك بدرجة الثقة ٩٥٪.

نحسب أولاً أحسن تقدير للانحراف المعيارى للمجتمع (ع) وهو يساوى:

$$\hat{\text{ع}} = \text{ع} \times \sqrt{\frac{\text{ن}}{\text{ن} - ١}} = ٥ \times \sqrt{\frac{٢٥}{٢٤}} = ٥,١$$

$$\text{ويكون الخطأ المعيارى للعينة} \quad \frac{\hat{\text{ع}}}{\sqrt{\text{ن}}} = \frac{٥,١}{\sqrt{٢٥}} = ١,٠٢$$

وحيث أن الانحراف المعيارى للعينة هو المعروف فتستخدم فى هذه الحالة قيمة (ت) المناظرة لدرجة الثقة ٩٥٪. ولدرجة الحرية (٢٥ − ١). ومن جدول (ت) يظهر أن هذه القيمة تساوى ٢,٠٦٤.

∴ متوسط انحدارات جمع روافد الرتبة الأولى = متوسط الانحدارات فى العينة ± قيمة (ت) × الخطأ المعيارى.

∴ متوسط انحدارات جميع الروافد = ٢٠ ± ٢,٠٦٤ × ١,٠٢

= ٢٠ ± ٢,١١

أى أن هناك احتمال مقداره ٩٥٪ أن يكون متوسط انحدارات جميع الروافد النهرية من الرتبة الأولى فى هذا الحوض النهرى عبارة عن قيمة تتراوح بين ١٧,٨٩ درجة و ٢٢,١١ درجة.

مثال (٥):

فى المثال رقم (٢) إذا رأت مصلحة الضرائب أن التكلفة المخصصة لفحص الإيرادات السنوية للمحلات الصغيرة فى محافظة الإسكندرية لاتكفى لدراسة عينة كبيرة، فسحبت عينة عشوائية حجمها ٢٦ محلاً، ولنفرض أن المتوسط الحسابى لإيرادات هذه المحلات هو أيضاً ١١٠٠ جنيه والانحراف المعيارى المحسوب من بيانات هذه العينة هو ١٢٠ جنيهاً أيضاً. والمطلوب إنشاء فترة ثقة بدرجة الثقة ٩٥٪.

بما أن حجم العينة صغيراً (ن أقل من ٣٠) والانحراف المعيارى للعينة معلوما، فإننا نستخدم قيمة (ت) بدرجات الحرية (ن - ١) = ٢٥ التى تجعل طرفى المنحنى تساوى ١ - ٩٥, = ٠٥, هى ٢,٠٦. والخطأ المعيارى لهذه العينة يحسب له قيمة أحسن تقدير للانحراف المعيارى للمجتمع وهى:

$$\hat{ع} = عـ \times \sqrt{\frac{ن}{ن - ١}}$$

$$= ١٢٠ \times \sqrt{\frac{٢٦}{٢٦ - ١}}$$

$= ١٢٠ \times ١,٠٢ = ١٢٢,٤$

الخطأ المعيارى $= \frac{\hat{ع}}{\sqrt{ن}} = \frac{١٢٢,٤}{\sqrt{٢٥}} = ٢٤,٤٨$

ويكون الحد الأدنى لفترة الثقة ٩٥٪ $= ١١٠٠ - ٢,٠٦ \times ٢٤,٤٨$

$= ١٠٥,١٧$ جنيهاً.

الحد الأعلى لفترة الثقة ٩٥٪ $= ١١٠٠ + ٢,٠٦ \times ٢٤,٢٨$

$= ١١٤٩,٣$ جنيهاً.

أى إذا كررنا هذه النسبة بعدد كبير جداً من المرات فإننا نتوقع أن معلمة المجتمع (مَ) تقع قيمتها بين ١١٤٩,٨٣، ١٠٥٠,١٧ جنيها فى ٩٥٪ من الحالات.

وبمقارنة حدى فترة الثقة السابقة بمثيلتها التى سبق تقديرها لهذا المثال فى حالة حجم العينة الكبير وباستخدام التوزيع المعتدل المعيارى نجد أن فترة الثقة للعينة الصغيرة أكبر من فترة الثقة للعينة الكبيرة لأن منحنى التوزيع (ت) أكثر تفرطحاً من منحنى التوزيع المعتدل وذلك لأنه يأخذ فى اعتباره الخطأ الناشئ من تقدير الانحراف المعيارى المحسوب من بيانات العينة الصغيرة. وبصفة عامة إذا كان الانحراف المعيارى للمجتمع (ع) معلوما فإن فترة الثقة الخاصة بعدد كبير من العينات من نفس الحجم لها مدى معين ثابت بالرغم من اختلاف قيمة مراكزها (المتوسطات الحسابية للعينات) وذلك لأن الخطأ المعيارى $\left[خ.م = \left(\frac{ع}{ن}\right) \right]$ مقدارها ثابت لكل العينات. أما إذا كانت قيمة الانحراف المعيارى للمجتمع (ع) غير معلومة فإن الخطأ المعيارى (خ.م) فى هذه الحالة لايكون مقدار ثابتاً بل سيختلف من عينة لأخرى وتبعاً لذلك فإن فترة الثقة للمتوسطات الحسابية المختلفة

المحسوبة من عينات ذات حجم متساوى لها مراكز مختلفة ومدى مختلف أيضاً. ويوضح ذلك الجدول التالى (جدول رقم: ١٠ - ١) والذى يعتمد على بيانات خاصة بالمثال رقم (٣).

جدول رقم (١٠ - ١)

العلاقة بين الانحراف المعيارى والخطأ المعيارى وفترة الثقة

ن	سَ	ع	غم	حدود الثقة ٩٥٪	فترة (مدى) الثقة ٩٥٪
أ- فى حالة معرفة قيمة (ع)					
١٠٠	٥٠	٢٠	٢,٠٠	٥٤,٠٠ ، ٤٦,٠٠	٨,٠٠
١٠٠	٦٠	٢٠	٢,٠٠	٦٤,٠٠ ، ٥٦,٠٠	٨,٠٠
١٠٠	٧٠	٢٠	٢,٠٠	٧٦,٠٠ ، ٦٤,٠٠	٨,٠٠
١٠٠	٨٠	٢٠	٢,٠٠	٨٨,٠٠ ، ٧٢,٠٠	٨,٠٠
ب- فى حالة عدم معرفة قيمة (ع)					
١٠٠	٥٠	١٠	١,٠٠	٥٢,٠٠ ، ٤٨,٠٠	٤,٠٠
١٠٠	٦٠	٢٠	٢,٠٠	٦٤,٠٠ ، ٥٦,٠٠	٨,٠٠
١٠٠	٧٠	٣٠	٣,٠٠	٧٦,٠٠ ، ٦٤,٠٠	١٢,٠٠
١٠٠	٨٠	٤٠	٤,٠٠	٨٨,٠٠ ، ٧٢,٠٠	١٦,٠٠

التقدير من نسبة العينة:

كثيراً ما تواجه الباحث الجغرافى بعض الحالات التى لايمكن فيها قياس المفردات المشاهدة ولكن يمكن حصر المفردات المشاهدة التى لها خاصية معينة، فمثلاً يمكن حصر العمال الإناث من الذكور فى صناعة النسيج أو حصر

المساحات المزروعة أذرة من المزروعات الصيفية على المستوى القومى. وعموماً فإن مفردات المجتمع يمكن أن تنقسم إلى أكثر من قسمين حسب طبقاً للصفات أو الخصائص المراد دراستها، كأن تنقسم مفردات المجتمع السكانى - حسب الحرف - إلى مفردات حرفتها الزراعة وثانية حرفتها الصناعة وأخرى حرفتها التجارة ... إلخ. ولكن فى بعض الأحيان، قد يتكون المجتمع من مجموعتين أو قسمين متميزين أحدهما له صفة أو خاصية معينة والآخر ليس فيه هذه الصفة أو الخاصية. وبعبارة أخرى يمكن أن نقسم مفردات المجتمع إلى وحدات موجبة وأخرى سالبة طبقاً للخاصية المراد إختبارها ودراستها، وتكون الوحدات الإيجابية هى الوحدات التى تتصف بهذه الخاصية بينما لاتتصف الوحدات السلبية بهذه الخاصية. فمثلاً يمكن تقسيم مجتمع الذكور فى سن معينة إلى أميين ومتعلمين، أو تقسيم مجتمع المواليد إلى أطفال ذكور وإناث، أو تقسيم مجتمع إنتاج إحدى الآلات إلى إنتاج معيب وآخر غير معيب ... إلخ. وقد يهمنا أحياناً أن نعرف نسبة كل مجموعة أو قسم (أى المفردات التى تمتلك الصفة المراد دراستها) فى المجتمع. ولكنه فى معظم الأحيان لايمكن قياس أو تحديد نسبة الصفة التى تتصف بها مفردات مجتمع ما فى المجتمع كله عن طريق الحصر الشامل لسبب من الأسباب التى شرحناها سابقاً، وعليه فإننا نقوم بسحب عينة عشوائية من هذا المجتمع ونحدد منها نسبة المفردات التى تمتلك الصفة التى نريد دراستها، وتؤخذ قيمة النسبة فى العينة كمقدر نقطة غير متحيز لنسبة المجتمع. فمثلاً إذا كانت نسبة الأمية المحددة من بيانات عينة مسحوبة عشوائياً من مجتمع محافظة ما هى ق = ٠,٥٥، فى هذه الحالة يمكن اعتبار أن نسبة الأمية فى هذه المحافظة كلها (د) هى ٠,٥٥ وهنا يمكن أن نعتبر أن ق هى تقدير نقطة غير متحيز للمعلمة (د)

وبصفة عامة إذا افترضنا أن نسبة عدد مفردات المجتمع التى تتصف بهذه الصفة هى (ق) وكانت النسبة (ق) قريبة جداً من الصفر أو الواحد الصحيح فإن التوزيع الاحتمالى للنسبة المحسوبة من البيانات المشاهدة فى عينة حجمها (ن) كبير

نسبياً ويقترب من التوزيع المعتدل الذى متوسطه الحسابى يساوى (د) وتباينه هو $\frac{\text{د}\,(١-\text{د})}{\text{ن}}$. ومعنى ذلك أن المتوسط الحسابى لقيم (ق) المحسوبة من كل العينات المختلفة (توزيع المعاينة للنسب) الممكنة المتساوية الحجم يساوى (د)، كما أن تباين توزيع (ق) يقل إذا كبر حجم العينة (ن). وبذلك إذا كانت (ن) كبيرة بدرجة كافية فإن قيم (ق) تتركز حول (د)، أى أن تشتتها حول المتوسط نادراً ما يكون بمقدار كبير.

ويمكن تقدير فترة النسبة فى مجتمع باستخدام عينة كبيرة الحجم (ن أكبر من ٣٠) عن طريق المعادلة الآتية:

تقدير فترة النسبة = $\text{ق} \pm \text{ز}_{\alpha\,١/٢} \times \text{خ.م}_{\text{ق}}$

حيث ق هى النسبة فى العينة، خ. مق هى الخطأ المعيارى للنسبة. وتكون بذلك حدود الثقة للنسبة فى المجتمع كما يلى:

حدود الثقة للنسبة = $\text{ق} \pm \text{ز}_{\alpha\,١/٢} \times \sqrt{\frac{\text{ق}\,(١-\text{ق})}{\text{ن}}}$ ... (١٠ - ٤)

وذلك فى حالة إذا كانت المعاينة من مجتمع محدود أو إذا كانت المعاينة بإرجاع من مجتمع محدود أما إذا كانت المعاينة بدون ارجاع من مجتمع محدود (حجمه N) فإن:

حدود الثقة بالنسبة فى المجتمع =

$$\text{ق} \pm \text{ز}_{\alpha\,١/٢} \times \sqrt{\frac{\text{ق}\,(١-\text{ق})}{\text{ن}}} \times \sqrt{\frac{N-\text{ن}}{N-١}} \quad ...(١٠-٥)$$

وتسمى القيمة $\sqrt{\frac{\text{ق}\,(١-\text{ق})}{\text{ن}}}$ بالخطأ المعيارى لتوزيع إحصائية نسبة العينة (ق).

مثال (٦):

سحبت عينة عشوائية من ٢٠٠ أسرة (حجم كل منها ٥ أفراد) من سكان منطقة معينة لمعرفة رأى هذه الأسر فى تطبيق أسلوب جديد لتنظيم النسل، فوجد أن ١٢٠ أسرة تستخدم الأسلوب المراد تطبيقه. قدر بدرجة الثقة ٩٥٪ نسبة الأسر المستخدمة للأسلوب الجديد لتنظيم النسل فى هذه المنطقة.

نسبة الأسر المستخدمة للأسلوب الجديد لتنظيم النسل (ق)

$$= \frac{\text{عدد الأسر}}{\text{حجم العينة}} = \frac{١٢٠}{٢٠٠} = ٠,٦٠$$

$$١ - ق = ١ - ,٦ = ,٤$$

$$\text{درجة الثقة} = ١٠٠ - \alpha = ٩٥٪$$

$$\alpha = ٥٪ \text{ ، } {}^{١}/_{٢}\, \alpha = ٢,٥٪$$

$$\therefore ز_{{}^{١}/_{٢}\alpha} = \pm ١,٩٦$$

الحد الأعلى لفترة الثقة ٩٥٪ للأسر المستخدمة للأسلوب الجديد لتنظيم النسل

$$= ,٦ + ١,٩٦ \times \sqrt{\frac{,٦ \times ,٤}{٢٠٠}}$$

$$= ,٦ + ,٠٦٨ = ٠,٦٦٨$$

الحد الأدنى لفترة الثقة ٩٥٪ للأسرة المستخدمة للأسلوب الجديد لتنظيم النسل

$$= ,٦ - ١,٩٦ \times \sqrt{\frac{,٦ \times ,٤}{٢٠٠}}$$

$$= ,٦ - ,٠٦٨ = ٠,٥٣٢$$

وعلى ذلك فإن نسبة الأسر المستخدمة للأسلوب الجديد لتنظيم النسل فى هذه المنطقة يقع بين (٠,٦٦٨، ٠,٥٣٢) وذلك بدرجة ثقة ٩٥٪ ونسبة خطأ ٥٪.

مثال (٧):

فى استطلاع للرأى العام بالعينة سحبت عينة عشوائية حجمها ١٠٠ من جميع الناخبين فى حى معين بإحدى المدن حيث دلت على أن أصوات ٥٥ ٪ منهم ستكون فى صالح مرشح معين، أوجد حدود الثقة ٩٥ ٪، ٩٩ ٪، ٩٩,٧٣ ٪ للنسبة بين جميع الناخبين المؤيدين لهذا المرشح. وما هو حجم العينة التى يجب أخذها من الناخبين بحيث يكون ٩٥ ٪، ٩٩,٧٣ ٪ منهم واثقين من أن هذا المرشح سوف يختار من مرشحين اثنين.

حدود الثقة ٩٥ ٪ لنسبة مجتمع الناخبين

$$= \text{ق} \pm \text{١,٩٦} \sqrt{\frac{\text{ق}\,(\text{١} - \text{ق})}{\text{ن}}}$$

$$= \text{٠,٥٥} \pm \text{١,٩٦} \sqrt{\frac{\text{٠,٥٥} \times \text{٠,٤٥}}{\text{١٠٠}}}$$

$$= \text{٠,٥٥} \pm \text{٠,١٠}$$

حدود الثقة ٩٩ ٪ لنسبة مجتمع الناخبين

$$= \text{٠,٥٥} + \text{٢,٥٨} \sqrt{\frac{\text{٠,٥٥} \times \text{٠,٤٥}}{\text{١٠٠}}}$$

$$= \text{,٥٥} \pm \text{٠,١٣}$$

حدود الثقة ٩٩,٧٣ ٪ لنسبة مجتمع الناخبين

$$= \text{,٥٥} \pm \text{٣} \sqrt{\frac{\text{٠,٥٥} \times \text{٠,٤٥}}{\text{١٠٠}}}$$

$$= \text{٠,٥٥} \pm \text{٠,١٥}$$

ويكون حجم العينة المطلوبة بدرجة الثقة ٩٥ ٪ هو:

$$\text{حجم العينة} = \text{ق} \pm \text{ز}_{\frac{1}{2}\alpha} \times \sqrt{\frac{\text{ق}\,(\text{١} - \text{ق})}{\text{ن}}}$$

$$= ,٥٥ \pm ز \; {}^{١}/_{٢} \; \alpha \sqrt{\frac{٠,٥٥ \times ٠,٤٥}{١٠٠}}$$

$$= ,٥٥ \pm \sqrt{\frac{\alpha \, ز \, ,٥٠}{ن}}$$

وحيث أننا استخدمنا التقدير ٠,٥٥ = ق على أساس البيانات السابقة، وبما أن المرشح سينجح فقط إذا حصل على أكثر من ٥٠٪ من أصوات مجتمع الناخبين، فإنه يجب أن تكون القيمة $\sqrt{\frac{\alpha \, ز \, ,٥٠}{ن}}$ أقل من ٠٥,

حجم العينة لدرجة الثقة ٩٥٪ هو: ٠,٥ = $\sqrt{\frac{\alpha \, ز \, ,٥٠}{ن}}$

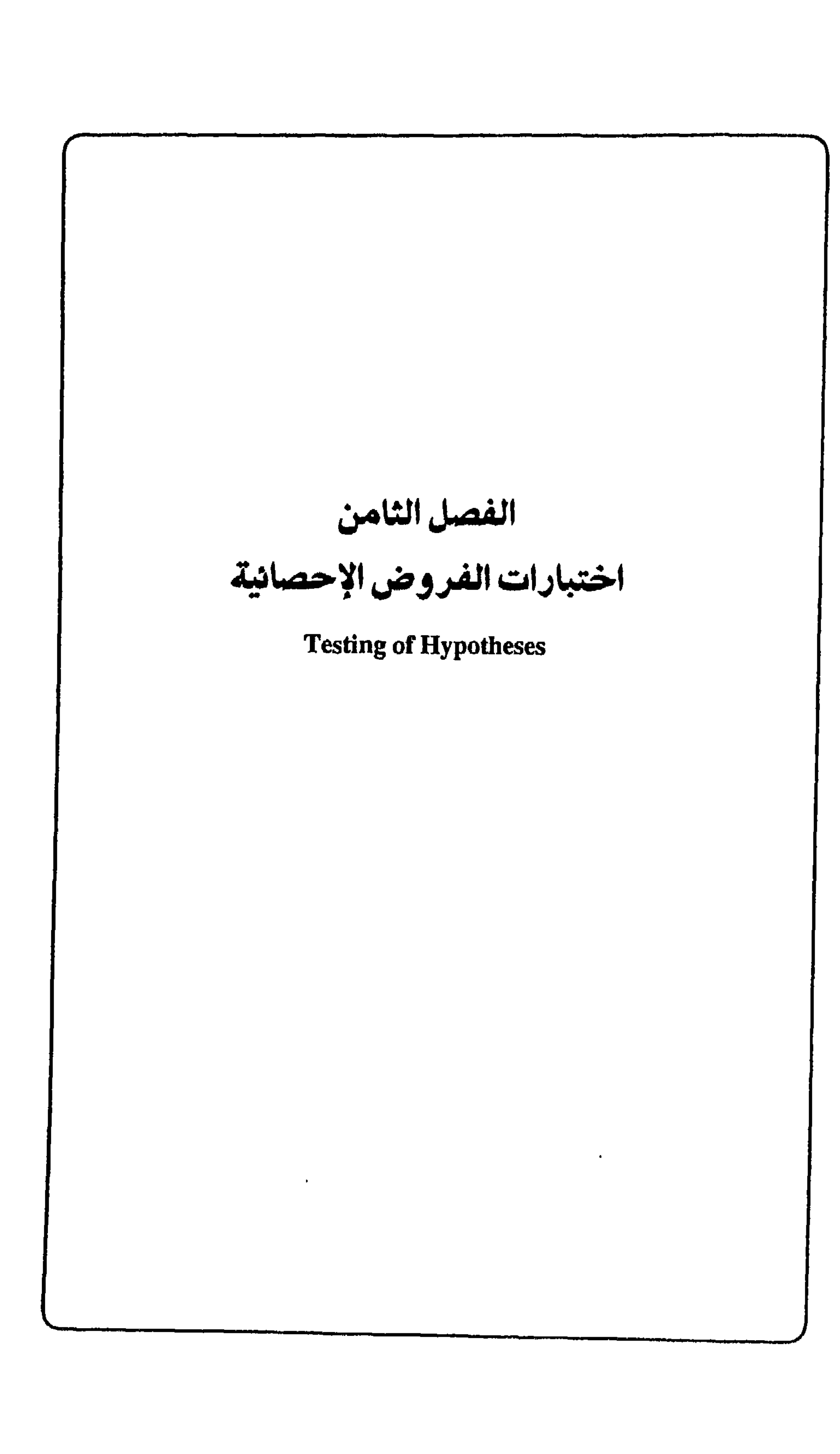
$$,٠٥ = \sqrt{\frac{١,٩٦ \times ,٥}{ن}}$$

ن = ٣٨٤,٢ (٣٨٥ ناخباً على الأقل)

حجم العينة لدرجة الثقة ٩٩,٧٣٪ هو: ٠٥, = $\sqrt{\frac{\alpha \, ز \times ,٥٠}{ن}}$

$$,٠٥ = \sqrt{\frac{٣ \times ,٥}{ن}}$$

ن = ٩٠٠ ناخباً على الأقل.



الفصل الثامن

اختبارات الفروض الإحصائية

Testing of Hypotheses



الفصل الثامن •

اختبارات الفروض الإحصائى

رأينا فى الفصل السابق كيف يمكن الاعتماد على توزيعات المعاينة لإيجاد فترة الثقة لبعض معالم المجتمع المجهولة. وفى هذا الفصل سندرس بعض اختبارات الفروض الإحصائية الهامة المبنية على أساس المتوسطات الحسابية والتباين للعينات وسنجد أن هناك صلة وثيقة بين التقدير الإحصائى والاختبار الإحصائى. وفى اختبارات الفروض الإحصائية تواجهنا مشكلة إتخاذ قرار بقبول فرض معين أو رفضه، ويتم اتخاذ هذا القرار بناء على البيانات التى تحصل عليها من عينة. فمثلاً إذا قلنا أن متوسط كمية الأمطار فى الاقليم (أ) يساوى متوسط كمية الأمطار فى القليم (ب) فإننا نطرح بذلك فرض يحتمل الصواب والخطأ: بمعنى أن هناك إحتمال أن يكون متوسط كمية الأمطار متشابهاً فى الاقليمين. ويتم اتخاذ قرار بقبول أو رفض هذا الفرض بعد أخذ عينة من كميات الأمطار فى فترة محددة وحساب متوسطهما للإقليمين، وذلك لأنه من الصعب كما عرفنا جمع البيانات عن مجتمع كميات الأمطار بأسلوب الحصر الشامل، أى بشكل دقيق، لذا فإنه يجب أن يكون اختيار العينة صحيحاً حتى تكون النتائج النهائية متشابهة إلى حد كبير النتائج التى يمكن الحصول عليها لو استخدمنا بيانات المجتمع كله. ولما كانت النتائج التى تستقى من إختبار أية عينة غير ممثلة تمثيلاً كاملاً أو غير مطابقة

تماماً لنتائج المجتمع فإن الفرض الإحصائى الخاص بمجتمع ما هو قول يحتمل الصواب والخطأ ولابد من جمع مجموعة من البيانات لمعرفة مدى انطباق صحة أو عدم صحة هذا الفرض على النتائج المتحصل عليها. فإذا كانت النتائج تتفق مع الفرض يقبل الفرض وبالتالى يمكن تعميمه. أما إذا لم تتفق النتائج مع الفروض يرفض الفرض. ويتم قبول أو رفض باستخدام الأساليب الإحصائية الكمية التى تتيح للباحث إتخاذ القرار المناسب فى ظل ظروف التشكك فى عدم التأكد.

ولاشك أن إختبارات الفروض واتخاذ قرار بشأنها يعد من أصعب الأمور. وللتسهيل فى عرض أسلوب التحليل الكمى ستتعرض فقط للمشكلات التى تشتمل على فرضين لاتخاذ قرار بتفضيل أحدهما على الآخر ، وذلك بعد تطبيق القواعد الرئيسية لاختبار هذه الفروض.

قواعد اختبار الفروض الإحصائية:

يمكن تحديد الأسس والقواعد اللازمة لإجراء إختبارات الفروض الإحصائية على النحو التالى:

١- وضع الفروض (فرض العدم والفرض البديل).

٢- تحديد مستوى المعنوية (مستوى الدلالة).

٣- تحديد التوزيع النظرى (الاحتمالى) للإحصائية المختبرة.

٤- إستخدام بيانات العينة لحساب قيمة إحصائية الاختبار وإستخدام التوزيع النظرى لإتخاذ القرار الإحصائى الخاص برفض أو قبول فرض العدم عن طريق تحديد المنطقة الحرجة (منطقة الرفض) ومنطقة القبول تحت منحنى التوزيع النظرى.

وفيمايلى مناقشة تفصلية لكل قاعدة من القواعد السابقة.

وضع الفروض:

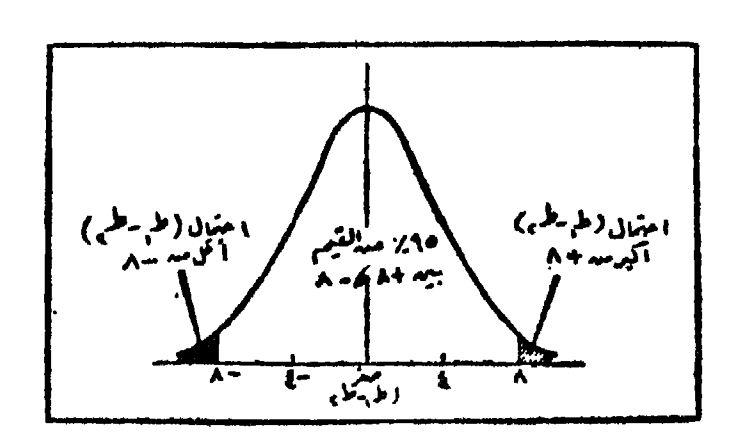
إن أولى الخطوات لإجراء اختبار الفرض هو التعبير عنه رياضياً (أى وضع افتراض معين للمعلمة المراد دراستها ثم يختبر هذا الافتراض فى مقابل المقياس المحسوب من البيانات المشاهدة من العينة) فإذا أردنا إختبار مدى تفوق الأقليم (أ) على الاقليم (ب) فى كمية الأمطار فإن الفرض المناسب فى هذه الحالة هو أن نفترض أن متوسطات كمية الأمطار متساوية فى الاقليمين أى أن متوسط كمية الأمطار فى الاقليم الأول (ط١) يساوى متوسط كمية الأمطار فى الأقليم الثانى (ط٢)، ويصبح الفرض المختبر Testing Hypothesis كمايلى:

الفرض المختبر: ط١ = ط٢

ويمكن التعبير عن هذا الفرض بصورة أخرى بأن نقول أن الفرق بين متوسطى كمية الأمطار فى الاقليمين يساوى صفرا، أى أن الفرض ينص على عدم وجود فرق بين المتوسطين (ط١ − ط٢ = صفر). ويسمى الفرض فى هذه الحالة بفرض العدم Null Hypothesis ويرمز له بالرمز (H_0)، وهو الفرض الذى لا يتفق مع البيانات المشاهدة.

فإذا قبل فرض العدم فإن ذلك يعنى أن النتائج جاءت مؤيدة له، أما إذا رفض الفرض فمعنى ذلك أن النتائج لم تكن مؤيدة له، ولذا فإننا نضطر إلى البحث عن الفرض البديل Alternative Hypothesis ويرمز له بالرمز (H_1) وفى المثال بين أيدينا يكون الفرض البديل هو أن متوسطى كمية الأمطار فى الاقليمين غير متساويين، أى أن ط١ = ط٢. ويعرف هذا النوع من الفرض البديل، الذى ينص على وجود فرق بين المتوسطين بالفرض ثنائى الطرف أو ثنائى الجهة - أى أنه فرض غير محدد Non- Directional . أما إذا توفر للباحث من الأدلة ما يجعله يعتقد بأنه فى حالة عدم تساوى المتوسطين فإن المتوسط الأول يتفوق على المتوسط

الثانى أو أن المتوسط الثانى يتفوق على المتوسط الأول، أى أن ط$_1$ › ط$_2$ أو ط$_1$ ‹ ط$_2$، فإن هذا الفرض يعرف بالفرض احادى الطرف أو أحادى الجهة – أنه فرض محدد Directional. وتتمثل الحالة الأولى فى النصف الأيمن من منحنى التوزيع العينى، بينما تتمثل الحالة الثانية فى النصف الأيسر من المنحنى كما نرى فى الشكل رقم (٨-١).



شكل رقم (٨-١)

لتوزيع العينى للاختلاف بين المتوسطات مقدراً بالقيم المعيارية

فمثلاً إذا كان الفرق أكبر من ٨ من الوحدات المعيارية فإنه لا يمكن منه تحديد ما إذ كانت ط$_1$ أكبر من أو أقل من ط$_2$ ولكن كما نرى أن هذا الفرق يتمثل (يقع) فى المساحة تحت طرفى التوزيع (المساحات المظللة فى الرسم) والتى تشتمل على ٥٪ (٠,٠٥) من احتمال تكرارات هذا الفرق والتى تتوزع على أساس ٢,٥٪ من المساحة الكلية تحت المنحنى (أو احتمال ٠,٢٥) فى الطرف الأيمن من منحنى التوزيع، ومثل هذا المقدار فى الطرف الأيسر. ومن الناحية الأخرى إذا كان الفرض البديل محدداً أى إذا كان الفرق موجباً (ط$_1$ › ط$_2$) أو سالباً (ط$_1$ ‹ط$_2$) أو سالباً (ط$_1$ ‹ ط$_2$)، فإنه يمكن تحديد نصف منحنى التوزيع

الذى يوافق هذا الفرق أو الذى تقع فيه قمة الفرض البديل. وكقاعدة عامة إذا كان الفرض البديل محدداً فإن اختبار هذا النوع من الفروض يسمى إختباراً من طرف واحد One-Taild Test، أما إذا كان الفرض البديل غير محدد فإن إختباره يسمى الاختبار ثنائى الطرف Two- tatild tests.

مما سبق يمكن أن نستنتج أن هناك أربع حالات لقبول أو رفض فرض العدم (أو رفض أو قبول الفرض البديل) هى كمايلى:

١- أن يكون فرض العدم صحيحاً وأن تؤيد نتائج اختبار العينة صحته، أى يقترب المقياس الاحصائى المحسوى من العينة من المقياس الاحصائى النظرى، وفى هذه الحالة يقبل فرض العدم ويكون قرار القبول صائباً.

٢- أن يكون فرض العدم صحيحاً ولكن لا تؤيد نتائج العينة صحته، فتكون المحصلة هى رفض هذا لفرض، وبذلك يكون هناك خطأ فى الحكم على فرض العدم برفضنا له بينما هو فى الواقع فرض صحيح وقبولنا للفرض البديل وهو فرض غير صحيح. ويعرف الخطأ فى هذه الحالة بالخطأ من النوع الأول Type I error ويرمز له بالرمز α، أى أنه احتمال رفض فرض العدم بالرغم من أنه فى الواقع صحيح.

٣- أن يكون فرض العدم غير صحيح بينما تأتى العينة بما يثبت ذلك، وتكون المحصلة هو قبول فرض العدم وبذلك يكون هو قبول فرض العدم وبذلك يكون هناك خطأ فى قرار القبول لفرض العدم وهو فى الواقع غير صحيح ورفض الفرض البديل وهو فرض صحيح، ويعرف هذا النوعين من الخطأ بالخطأ من النوع الثانى Type II Error، ويرمز له بالرمز B أى أنه احتمال قبول فرض العدم بالرغم من أنه فى الواقع غير صحيح.

٤- أن يكوت فرض العدم غير صحيح ولكن لا تأتى نتائج العينة بما يثبت ذلك، وتكون المحصلة هى رفض فرص العدم وهو فى الواقع غير صحيح، وقبول الفرض البديل وهو فى الواقع صحيح، وبذلك يكون القرار برفض فرض العدم فى هذه الحالة سليماً.

وعليه يتضح لنا أنه عند قبول أو رفض فرض العدم فإننا نتعرض لنوعين من الأخطاء. ويمكن تلخيص القرارات الممكنة السابقة والأخطاء الناجمة عنها في الجدولين التاليين:

جدول رقم (٨-١)

حالات قبول أو رفض فرض العدم

النتيجة	القرارات الممكنة	نوع الفرض: الفرض البديل	نوع الفرض: فرض العدم
القرار صائب	قبول فرض العدم	غير صحيح	صحيح
القرار خاطئ	رفض فرض العدم	غير صحيح	صحيح
القرار خاطئ	قبول فرض العدم	صحيح	غير صحيح
القرار صائب	رفض فرض العدم	صحيح	غير صحيح

جدول رقم (٨-٢)

أنواع أخطاء القرارات رفض أو قبول فرض العدم

نوع الفرض	الواقع	
فرض العدم	فرض العدم صحيح	فرض العدم غير صحيح
رفض	خطأ من النوع الأول (α) Type I Error	صائب (β)
قبول	صائب (١-α)	خطأ من النوع الثاني (β) Type II Error

اختبار المعنوية (الدلالة) Test of Significance

يعتمد تحديد قيم التوزيعات النظرية لاحصائيات العينات (المعايير الاحصائية) على نسبة أو احتمال الخطأ المسموح به لقبول أو رفض الفروض الاحصائية. وتعرف نسبة الخطأ أو الاحتمال المسموح به بمستوى المعنوية أو مستوى الدلالة

Level of Significance الذى يكون اختباره فى الواقع الخطوة التالية على طريق إختبارات الفروض الإحصائية. وعند تحديد مستوى المعنوية يجب الأخذ فى الاعتبار نوعى الخطأ فى رفض العدم وهو صحيح، وقد عرفنا ذلك بالخطأ من النوع الأول. فمثلاً إذا قررنا قبول حدوث خطأ من النوع الأول فى خمس مرات كل مائة مرة فإن قرارنا هذا يعنى أنه فى هدد كبير من التجارب نتوقع أن نرفض فرض العدم وهو فى الواقع صحيح ٥٪ من المرات وبذلك يكون الحد الأقصى الذى قررنا قبوله لاحتمال (a) وتسمى قيمة الاحتمال (a) بمستوى المعنوية أو مستوى الدلالة. من هذا نرى أن مستوى المعنوية هو احتمال حدوث خطأ من النوع الأول، أى احتمال رفض فرض العد م وهو فى الواقع صحيح. وبالتالى فإن القيمة التى نحددها لهذا الاحتمال تعتبر الأساس فى الحكم على وجود فروق جوهرية، من الناحية الإحصائية Statistically Significance، بين إحصائيات العينات وبين معالم المجتمع أو ارجاع هذه الفروق إلى الصدفة. وقد جرت العادة على اختبار α = ٠,٠٥ أو ٠,٠١ فإذا كانت a = ٠,٠٥ ورفضنا فرض العدم فإننا نستنتج أن نتيجة العينة تختلف جوهرياً عن فرض العدم بمستوى معنوية ٠,٠٥ ومن الناحية الأخرى قد تؤدى نتائج العينة إلى قبول فرض العدم وهو فى الواقع غير صحيح، فنكون قد وقعنا فى خطأ من النوع الثانى، ويرمز لاحتمال حدوث هذا الخطأ بالرمز (β). وبالتالى فإن إحتمال رفض فرض العدم وهو غير صحيح يساوى (1- β) ويسمى هذا الاحتمال بقوة الاختبار Power of the Test وتسمى قيمة الاحتمال (β). بمستوى الثقة Level of Confidence، وهو يعكس مستوى الدلالة حيث يشير إلى الخطأ فى قبول فرض العدم وهو غير صحيح. وقد جرت العادة على اختبار β = ٠,٩٥ أو ٠,٩٩. فمثلاً إذا كانت β= ٠,٩٥ وقبلنا فرض العدم فإن قرارنا هذا يعنى أنه إذا تكررت التجربة عدد كبير من المرات نتوقع أن تقبل فرض العدم وهو فى الواقع يغر صحيح ٠,٩٥ من المرات.

وتجدر الإشارة هنا إلى أن قيمة (α) يتم تحديدها بافتراض صحة فرض العدم بينما تحسب قيمة (β) بافتراض صحة الفرض البديل. كما أنه من الممكن تقليل احتمال الخطأ β ولكن ذلك يكون على سحاب زيادة احتمال الخطأ (α)، أما إذا أردنا تقليل احتمال الخطأ (β) بينما يظل احتمال الخطأ (α) ثابت فإنه يجب زيادة حجم العينة، فكلما زاد حجم العينة كلما انخفضت قيمة الانحراف المعيارى

وأصبحت خصائص العينة أكثر تمثيلاً لمعالم المجتمع الذى سحبت منه. ونظراً لأنه قد جرت العادة على تحديد مستوى المعنوية (α) قبل إجراء الاختبار فإنه يمكن التحكم فى الخطأ β، حيث أن قيمة (a) متممة لقيمة (β) ،ولكن ذلك يرتبط بمدى خطورة أو أهمية النتائج المترتبة على الاختبار. فكلما كانت النتائج المترتبة على الوقوع فى الخطأ (α) أكثر خطورة من مثيلتها المترتبة على الوقوع فى الخطأ (a) كلما قلت قيمة (α) التى يختارها الباحث، مثل ٠,٠١ أو ٠,٠٠١، وبالتالى تزداد قيمة β، أى أن β= (١- ٠,٠١ = ٠,٩٩) أو β = (١- ٠,٠٠١ = ٠,٩٩٩) ، أى يزداد احتمال الوقوع فى الخطأ من النوع الثانى مما يؤدى إلى نقص قوة الاختبار (١ - β) وتتمثل هذه الحالة فى أغلب التجارب المعملية. والعكس كلما كانت النتائج المترتبة على الوقوع فى الخطأ (β) أكثر خطورة من الوقوع فى الخطأ (α) . مثل ٠,٠٥ أو، ٠,١٠ ، وبالتالى تقل قيمة (β) أى أن β = (١- ٠,٠٥ = ٠,٩٥) أو β = (١- ٠,١٠ = ٠,٩٠)، أى ينقص احتمال الوقوع فى الخطأ من النوع الثانى مما يؤدى إلى زيادة قوة الاختبار. وتتمثل مثل هذه الحالة فى اختبار الفروض المتعلقة بأغلب المشاكل فى الجغرافية الاقتصادية.

تحديد التوزيع النظرى (الاحتمالى) للإحصائية المختبرة:

يعتمد قبول أو رفض الفروض الاحصائية، أو بمعنى آخر الاستدلال على صحة أو خطأ الفروض، على حساب بعض المقاييس الاحصائية من العينة أو العينات (والتى تعرف باحصائية الاختبار) ومقارنة هذه المقاييس بتلك المقاييس الاحصائية النظرية (أو مايعرف بالمعايير الإحصائية) والتى عن طريقها يمكن تقدير الخطأ فى قبول أو رفض الفرض الاحصائى. فإذا كانت المقاييس الأولى تقترب من الثانية فإنه يتم قبول الفرض المختبر والعكس صحيح ويمكن اختبار القيمة المعيارية للإحصائية والتى هى، كما ذكرنا، عباره عن الفرق بين الاحصائية المحسوبة من العينة ومعلمة المجتمع مقسوماً على الخطأ المعيارى. فمثلاً إذا كان لدينا توزيعاً عينياً يمكن حساب المتوسط الحسابى منه ووضعه فى صورة معيارية فإن قيمة المتغير المعيارى المحسوب يمكن مقارنتها بتوزيعها النظرى، وبالتالى يمكن تقرير امكانية قبول أو رفض فرض العدم.

وتحسب قيم إحصائية الاختبار (ز) فى حالة توفر بيانات عن امجتمع (أى فى حالة التوزيع المعتدل)، بينما تحسب قيمة (ت)، وقيمة (ف) وقيمة مربع كأى فى حالة العينات. والذى يحدد قرب قيم إحصائيات الاختبار (أى قبول أو رفض الفرض الاحصائى) هو التوزيع النظرى (الاحتمالى) لهذه الاحصائيات أى توزيع (ز) Z-Scores، توزيع (ت) T-distribution، توزيع (ف) F-distribution وتوزيع مربع كاى $\chi 2$- distribution. وهذه التوزيعات مدونة فى جداول خاصة لرجوع إليها عند تحديد المقارنة بين الفروض النظرية والفروض الحقيقية (أنظر ملاحق الجداول الاحصائية فى نهاية هذا الكتاب).

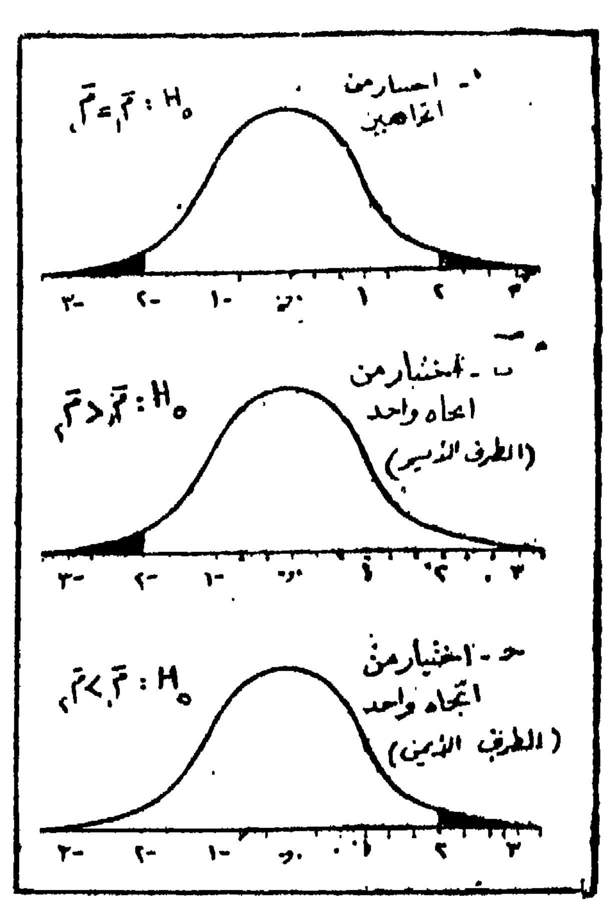
تحديد المنطقة الحرجة:

بعد تحديد قيمة مستوى المعنوية وبعد وضع إحصائية العينة فى الصورة المعيارية وتحديد التوزيع النظرى (الاحتمالى) لها يمكن تعيين منطقة الوقوع فى الخطأ من النوع الأول أو ما يسمى بالمنطقة الحرجة Critical Region أو منطقة الرفض Re-jection Region، وتحتوى هذه المنطقة من منحنى التوزيع على جميع القيم الحرجة التى تدعونا إلى رفض فرض العدم وهو صحيح، وذلك لأن إحتمال أن تقع نتيجة العينة فى هذه المنطقة إذا كان فرض العدم صحيح يساوى مستوى معنوية (دلالة) a.

وقد تمتد المنطقة الحرجة على طرفى منحنى التوزيع النظرى (الاحتمالى) أو قد تمتد على طرف واحد من طرفى التوزيع على حسب التجربة المراد إختبارها. ويسمى الاختبار فى الحالة الأولى بالاختبار ثنائى الجهة (الطرف) Two- tailed Test الذى يستخدم عندما يراد اختبار ما إذا كانت إحصائية عينة تختلف اختلافاً جوهرياً عن توقع المجتمع المفروض، وفى الحالة الثانية يسمى بالاختبار أحادى الجهة (الطرف) One- tailed Test الذى يستخدم عندما يراد اختبار الانحراف الموجب فقط أو الانحراف السالب فقط للإحصائية المحسوبة من بيانات العينة عن

متوسط المجتمع الحقيقى. فمثلاً إذا كانت الاحصائية لها توزيع معتدل وكان مستوى المعنوية a = ٠,٠٥ وقررنا إجراء اختبار ثنائى الطرف فإن المنطقة الحرجة تشتمل فى هذه الحالة على ٢,٥ ٪ من المساحة تحت المنحنى على كل طرف من طرفيه، وتسمى المنطقة الواقعة بين المنطقة الحرجة على الطرفين بمنطقة القبول Region of Acceptance وتمثل ٩٥٪ من المساحة تحت المنحنى. أما فى حالة الاختبار أحادى الطرف بمستوى معنوية = ٠,٠٥ نجد أن المنطقة الحرجة على هذا الطرف تشتمل على ٥ ٪ من المساحة الكلية، فإذا كانت المنطقة الحرجة على الطرف الأيمن فإن القيم الحرجة تقع على يمين قيمة إحصائية الاختبار المعيارية، وإذا كانت المنطقة الحرجة على الطرف الأيسر فإن القيم الحرجة تقع على يسار قيمة إحصائية الاختبار المعيارية وبذلك يكون احتمال رفض فرض العدم وهو فى الواقع صحيح أكبر فى الاختبار أحادى الطرف منه فى الاختبار ثنائى الطرف شكل (٨-٢).

ويتخذ القرار الاحصائى على أساس مقارنة النتائج المشاهدة من بيانات العينة بالنتائج النظرية، فإذا وقعت قيمة إحصائية الاختبار المحسوبة فى المنطقة الحرجة يرفض فرض العدم ويقبل الفرض لبديل، ويدل ذلك على وجود فرق جوهرى أو حقيقى بين الاحصائية والقيمة المفترضة لمعلمة المجتمع. أما إذا وقعت قيمة إحصائية الاختبار فى منطقة القبول نعتبر الفرق بين النتيجة المشاهدة والمفروضة للمجتمع هو فرق غير جوهرى أى أنه فرق ظاهرى ربما يرجع إلى الصدفة المطلقة لخطأ المعاينة.



شكل رقم (٨-٢) المنطقة الحرجة واختبارات الفروض الإحصائية

وسنقوم فى ما يلى بتطبيق الأسس والمبادئ لإجراء بعض اختبارات الفروض الإحصائية الخاصة بتوزيع المجتمع (التوزيع المعتدل) وذلك عن طريق حساب احصائية الاختبار (ز).

اختبار انتماء عينة لمجتمع متوسطة معلوم:

اذا سحبنا عشوائياً عدة عينات حجمها صغير (ن ‹ ٣٠) من مجتمع معتدل التوزيع متوسطه الحسابى (مَ) غير معلوم، فان التوزيع العينى يتبع توزيع المجتمع ويكون متوسطه الحسابى (سَ) مناظر لمتوسط المجتمع. أما إذا سحبنا عشوائياً عدة عينات حجمها كبير (ن › ٣٠) من مجتمع يتصف بقربه من الاعتدالية فى

التوزيع فان متوسط التوزيع العينى (سَ) يقترب من التوزيع المعتدل. وعليه فإنه كلما كان حجم العينة كبيراً كلما أدى ذلك إلى اعتدالية التوزيع العينى بغض النظر عن حالة توزيع المجتمع الذى سحبت منه هذه العينات.

ولاختبار الفرق بين متوسط عينة (سَ) ومتوسط مجتمع (مَ) تباينه ($ع^2$) معلوم فاننا نحسب احصائية الاختبار التى تساوى فى هذه الحالة:

$$(\text{ز}) = \frac{\frac{\bar{\text{س}} - \bar{\text{م}}}{\text{ع}}}{\sqrt{\frac{\text{ع}}{\text{ن}}}}$$

وفى حالة عدم معرفة تباين المجتمع ($ع^2$) فاننا نستخدم تباين العينة ($عـ^2$) بدلاً منه، وفى هذه الحالة تتبع احصائية الاختبار قيم (ت) المعيارية. وبازدياد حجم العينة (ن › ٣٠) فان التوزيع الأخير يقترب من التوزيع المعتدل، وتأخذ احصائية الاختبار الشكل التالى:

$$(\text{ز}) = \frac{\frac{\bar{\text{س}} - \bar{\text{م}}}{\text{ع}}}{\sqrt{\frac{\text{عـ}}{\text{ن}}}}$$

ولتوضيح ماسبق ذكره نعطى المثال الآتى:

مثال: أراد باحث دراسة النشاط التجارى للصيدليات فى مدينة الاسكندرية مستخدماً لذلك معياراً يتمثل فى حجم مبيعاتها اليومية بالجنيه. فلما سحب الباحث عينة عشوائية مكونة من ١٤٤ صيدلية وجد أن متوسط مبيعاتها ٩٩٨,٠ جنيه، فإذا كان الانحراف المعيارى لكل الصيدليات هو ٢٠ جنيهاً فهل يعنى ذلك أن متوسط مبيعات كل الصيدليات فى الاسكندرية هو ١٠٠٠ جنيه فى اليوم وذلك عند مستوى معنوية ٠,٠٥.

لإجراء الاختبار فى المثال السابق فاننا نتبع الخطوات التالية:

١- تحديد فرض العدم والفرض البديل:

مَ = ١٠٠٠ فى مقابل مَ ≠ ١٠٠٠

٢- تحديد مستوى المعنوية أو الدلالة (أى تحديد احتمال رفض فرض العدم، أو احتمال قبول الفرض البديل، وفى المثال قيمة α = ٠,٠٥)

٣- تحديد إحصائية الاختبار (ز).

٤- إستخدام بيانات العينة لاصدار القرار الإحصائى برفض فرض العدم أو قبول الفرض البديل (وهو فى المثال السابق من النوع ثنائى الطرف). وباستخدام جداول المنحنى المعتدل المعيارى نجد أن فرض العدم سيرفض إذا كانت قيمة (ز) المحسوبة أكبر من (أو مساوية) لقيمة (ز) النظرية = ± ١,٩٦.

وبإستخدام بيانات العينة السابقة نجد أن:

$$\text{الخطأ المعيارى} = \frac{\text{الإنحراف المعيارى}}{\sqrt{\text{ن}}} = \frac{٢٠}{\sqrt{١٤٤}}$$

$$= \frac{٢٠}{١٢} = ١,٦٦$$

وبحساب قيمة (ز) نجد أن:

$$(\text{ز}) = \frac{\bar{\text{س}} - \bar{\text{م}}}{\frac{\text{ع}}{\sqrt{\text{ن}}}} = \frac{٩٩٨,٠ - ١٠٠٠}{١,٦٦} = \frac{-٢}{١,٦٦} = -١,٢٠٤$$

وبمقارنة قيمة (ز) المحسوبة وهى -١,٢٠٤ فى المثال بالقيمة المستخرجة من جداول المنحنى المعتدل المعيارى عند مستوى معنوية ٠,٠٥ وهى -١,٩٦ نجد أن قيمة (ز) المحسوبة لاتقع فى منطقة الرفض، وبناء على ذلك نقبل فرض العدم

القائل بأن متوسط مبيعات الصيدليات فى الاسكندرية ١٠٠٠ جنيه فى اليوم الواحد.

مثال (٢):

فى المثال السابق إذا كان متوسط مبيعات الصيدليات فى العينة هو ٩٩٥ جنيه فى اليوم والمطلوب اختبار الفرض القائل بأن هذا المتوسط أكبر من أو يساوى المتوسط العام لجميع الصيدليات ١٠٠٠ جنيه فى اليوم فى مقابل الفرض البديل بأن المتوسط فى العينة أقل من ١٠٠٠ جنيه وذلك عند مستوى معنوية ٠,٠٥. باتباع نفس الخطوات السابقة يلاحظ أن:

H_o: سَ ≥ ١٠٠٠ H_1: سَ < ١٠٠٠

ويلاحظ أن الخطأ فى الفرض البديل هو خطأ من النوع الثانى β لاختبار أحادى الطرف. أى أن فرض العدم سيرفض عندما تكون س أصغر من ١٠٠٠ جنيه بدرجه المعنوية المذكورة. ومنطقة الرفض ستقع بالتالى على الطرف الأيسر للمنحنى المعتدل المعيارى شكل (٨-٢ب) وبحساب قيمة (ز) نجد أن:

$$(ز) = \frac{٩٩٥ - ١٠٠٠}{١,٦٦} = \frac{-٥}{١,٦٦} = -٣,٠١$$

وبمقارنة قيمة (ز) المحسوبة من العينة (-٣,٠١) بالقيمة المستخرجة من الجدول وهى ١,٦٤٠ نجد قيمة (ز) المحسوبة أقل من قيمة (ز) بادول ، أى أنها تقع فى منطقة الرفض. وبذلك نستطيع رفض فرض العدم للقائل بأن متوسط مبيعات الصيدليات فى العينة أكبر من أو يساوى المتوسط العام وهو ١٠٠٠ جنيه وذلك عند مستوى معنوية ٠,٠٥ أو بمعنى آخر قبول الفرض البديل القائل بأن متوسط مبيعات الصيدليات فى العينة أقل من ١٠٠٠ جنيه عند نفس مستوى المعنوية أو الدلالة.

مثال (٣):

في المثال الأول إذا كان متوسط مبيعات الصيدليات هو ١٠٠٥ جنيه في اليوم الواحد والمطلوب اختبار الفرض القائل أن متوسطات المبيعات في العينة أقل أو يساوى ١٠٠٠ جنيه فى مقابل الفرض البديل بأن متوسط المبيعات للعينة أكبر من ١٠٠٠ جنيه في اليوم عند مستوى دلالة أو معنوية ٠,٠٥. في هذا المثال ينحصر الاختبار في:

$$H_0: \bar{س} \leq ١٠٠٠ \qquad \text{فى مقابل } H_1: \bar{س} > ١٠٠٠$$

وفي هذا المثال نجد الفرض لابديل عكس نفس الفرض في المثال الثاني أو بمعنى آخر نجد أن الفرض البديل ذو طرف أيمن أى أنه يمكن رفض فرض العدم عندما تكون $\bar{س}$ أكبر من ١٠٠٠ جنيه بدرجة معنوية ٠,٠٥ (شكل ٨- ٢جـ). وبحساب قيمة (ز) نجد أن:

$$(ز) = \frac{\bar{س} - \bar{م}}{\frac{ع}{\sqrt{ن}}} = \frac{١٠٠٥ - ١٠٠٠}{١,٦٦} = \frac{+٥}{١,٦٦} = +٣,٠١$$

وبمقارنة قيمة (ز) المحسوبة وهى + ٣,٠١، بالقيمة النظرية فى جداول المنحنى المعتدل المعيارى عند مستوى معنوية ٠,٠٥ وهى + ١,٦٤٥، نجد أن قيمة (ز) المحسوبة أكبر من ١,٦٤ أى قيمة (ز) المحسوبة من العينة تقع في منطقة الرفض، لذلك نرفض فرض العدم القائل بأن متوسط المبيعات فى العينة أقل من المتوسط العام للمبيعات لكل الصيدليات فيى الاسكندرية عند مستوى معنوية ٠,٠٥، وقبول الفرض البديل الذى يقول أن متوسط العينة أكبر من المتوسط العام عند نفس مستوى الدلالة أو المعنوية.

اختيار الاختبارات الإحصائية:

يعتمد إختيار الاختبارات الإحصائية Statistical Tests على طبيعة وخصائص البيانات Characteristics of the data، والقيمة الفعلية للأساليب الكمية المرتبطة بها والمستخدمة فى عمليات البحث والتحليل، والافترضات الإحصائية عن المجتمعات التى تستقى منها البيانات. وكما ذكرنا فى مقدمة هذا الكتاب أن هناك ثلاثة أنواع من البيانات، حسب الطرق المختلفة التى تقاس بواسطتها، وهى البيانات الرسمية (النوعية) أو الوصفية Nominal data والبيانات الترتيبية Ordinal or Rank-ing data وبيانات الفترة Interval data. وكما عرفنا أن النوع الأول من البيانات يتصف بأنه قائم على أساس التعداد أو العد Counts بينما تكون القياسات Mes-surements أهم صفات النوعين الآخرين. كذلك قد تكون البيانات ذات قيم فردية أو ثنائية (أى مزدوجة) على أساس أن العد أو القياس فى مجموعة بيانات يماثل نظيرة فى المجموعات الأخرى – بالإضافة إلى أننا قد نكون بصدد مقارنة بيانات فعلية (حقيقية) لمجموعة واحدة، أو لمجموعتين أو أكثر بيانات توزيع نظرى. لكل ذلك فإن أنواع الاختبارات الإحصائية المستخدمة فى البحوث الجغرافية تختلف حسب نوعية البيانات والطرق التى قيست بها. فهناك إختبارات إحصائية لا تصلح أو لا يمكن تطبيقها إلا فى حالات بيانات الفترة بل أن معظم الأساليب تفترض أن البيانات المتوفرة هى من هذا النوع، أما البيانات الإسمية (النوعية) والترتيبية فلا يستخدم معهما إلا الأساليب الإحصائية البسيطة. ويوضح الجدول التالى (جدول رقم: ٨-٢) بعضا من الشروط التى يجب أن يلم بها الباحث قبل اختيار، كما يبين الأنواع المختلفة من الاختبارات الإحصائية وملائمة كل منها لخصائص وطبيعة البيانات.

جدول رقم (٨-٢)

خصائص البيانات وأنواع الاختبارات الإحصائية

المستخدمة في مقارنة الاختلاف بين قيم التوزيعات

	بيانات خام	نوع الجدولة	نوع الاختبار
مقارنة مجموعتين من البيانات	١- بيانات قائمة على العد ينتج عنها تكرارات إسمية: أ- في فئتين ب- في أكثر من فئتين	 قيم عددية قيم عددية	 مربع كاى مربع كاى
	٢- قياسات فردية من نوع بيانات الفترة	تكرارات رتب قيم عددية	مربع كاى أو كولمرجورروف سمير نوف مان - هوتيني (اختبار «ى») ستيودنت (أختبار «ت»)
	٣- قياسات ثنائية من نوع بيانات الفترة	رتب قيم عددية	ويلكوكسون ستيودنت (اختبار «ت»)
مقارنة أكثر من مجموعتين من البيانات	١- بيانات قائمة على العد ينتج عنها تكرارات إسمية	قيم عددية	مربع كاى
	٢- قياسات من نوع بيانات الفترة (فردية وثنائية غير متكافئة في عدد مفرداته)	تكرارات رتب قيم عددية	مربع كاى كروسكال- واليس (اختبار «هـ») تحليل التباين (اختبار «ف»)

أما من حيث القيمة الفعلية التى يتوقف عليها اختيار الاختبارات الإحصائية فنقصد بها القوة Power أو القدرة على التمييز بين الفرض الحقيقى والفرض غير الحقيقى، ويعتمد ذلك على حجم لعينات المختبرة من ناحية وعلى مدى فاعلية الاختبار نفسه من ناحية أخرى. فمثلاً إذا إخترنا أسلوباً بسيطاً فى البحث والتحليل فإن ذلك يتطلب سحب عينات كبيرة الحجم حتى يتحقق نفس مستوى القوة أو التمييز الذى تتصف به الأساليب ذات المقدرة والفاعلية. «وعليه فإننا نتوقع أن تكون هناك مقاييساً أكثر قوة وفاعلية يمكن استخدامها فى التحليل الاحصائى أكثر من غيرها حتى إذا كانت العينات المختبرة صغيرة الحجم. فمثلاً إذا كان حجم العينة المطلوب لتحقيق قوة أحد الاختبارات ذات الفاعلية فى التحليل هو ن١، وكان حجم العينة لتحقيق مستوى نفس القوة لاختبار آخر أقل فاعلية هو ن٢، فإن قوة الاختبار الأخير تساوى (ن١ ÷ ن٢) × ١٠٠٪، وتبعاً لذلك فإن إختبار له قوة تساوى ٨٠٪ (٤/٥) سوف يتطلب حجم عينة مقدارة ١٢٥٪ (٥/٤) من حجم العينة التى يتطلبها أكبر الاختبارات المتاحة قوة وأكثرها فاعلية.

أما عن الافتراضات الإحصائية عن المجتمعات ونوع توزيعاتها التكرارية فتبدو أهميتها فى أنها تتخذ كأساس لتقسيم «مجتمع» الاختبارات الإحصائية المستخدمة فى المقارنة، إلى «أسرتين» هما: (١) الاختبارات الكلاسيكية (القديمة) Classical أو البارامتيرية (المعلمية) Parametric tects، التى سادت كأسلوب عمل فى النظرية الإحصائية والممارسة العملية حتى وقت ليس بعيد، والتى يمكن تطبيقها فى حالات بيانات الفترة التى هى أكثر شيوعاً لأنها أكثر توافقاً مع افتراضات هذه الاختبارات عن المجتمعات. (٢) الاختبارات الحديثة أو اللابارامترية (غير اللامعلمية) Non- Parametric التى أكتسبت أهمية خاصة منذ الحرب العالمية الثانية ٣٩- ١٩٤٥ إذ أنه يمكن تطبيقها على البيانات الإسمية (النوعية) والترتيبية وبيانات الفترة على حد سواء. وتعد اختبارات النوع الأول أكثر قوة من إختبارات النوع الثانى، ولأنها تفترض أن مفردات المجتمع تتوزع توزيعاً معتدلاً كما أنها تستلزم فى حالة استخدام عينات صغيرة الحجم فى التحليل أن يكون التوزيع

المعتدل للمجتمع الذى سحبت منه هذه العينات مؤكداً، بينما لا تتقيد بذلك الاختبارات اللابارامترية.

وفى الحالات التى لا يكون فيها توزيع بيانات المجتمع معتدلاً يتم تحويل البيانات بطرق مختلفة، سبق شرحها، ليصبح توزيعها معتدلاً، حتى يمكن تطبيق الاختبارات البارميترية عليها، فإن لم يتحقق ذلك فيجب تطبيق الاختبارات اللابارامترية لأنها لا تشترط توزيعاً معتدلاً للبيانات. كما تبرز أهمية مثل هذه الاختبارات فى حالة إذا كانت العينات قيد الاختبار صغيرة الحجم، وهو ما سنناقشة بالتفصيل فى الفصلين التاليين.

الفصل التاسع
أساليب المقارنة البارامترية
(المعلمية)

أساليب المقارنة البارامترية (المعلمية)

تشترط الاختبارات البارامترية Parametric Tests والأساليب الكمية المرتبطة بها – لمقارنة معالم المجتمعات أو إحصائيات العينات – توفر الخصائص التالية فى بيانات المجتمع Fopulation قيد الفحص:

١– أن يكون توزيع البيانات توزيعاً معتدلاً (متماثلا)، أى أن معامل التوائه يساوى صفراً.

٢– أن تكون المفردات المشاهدة أو الحالات (Observations or Cases) مستقلة عن بعضها البعض، أى أن اختيار إحدى المفردات لايمنع إمكانية إختيار أى مفردة أخرى من المفردات المطلوب دراستها.

٣– أن يكون للمجتمعات المقارنة مع بعضها البعض تباينا Variance متساويا، أو بمعنى آخر أن يكون هناك تجانسا بين المجتمعات موضع المقارنة.

٤– أن تكون البيانات المقاسة والتى تجرى عليها الاختبارات من نوع بيانات الفترة Interval Data .

فإذا لم تتوفر هذه الشروط أو الخصائص فى البيانات، فإن تطبيق الاختبارات البارامترية عليها يكون غير مناسب وبالتالى تكون النتائج النهائية مضللة، ولهذا نلجأ إلى النوع الآخر من الاختبارات: لاختبارات غير البارامترية – وهى الاختبارات التى سنعرضها وسنناقش الأساليب المرتبطة بها فى الفصل التالى مباشرة (الفصل العاشر).

وتتم عملية مقارنة وتحليل البيانات المعتدلة التوزيع، لاختبار الفروق بين معالم

المجتمع وبين احصائية (المتوسطات الحسابي أو الانحراف المعياري) عينتين أو أكثر لمتغير واحد، ومعايرة نتائجها بواسطة عدة اختبارات بارامترية أكثرها شيوعاً هو اختبار ستيودنت (ت) Student-t test واختبار تحليل التباين (أو نسبة «ف») Analysis of Variance (F ratio test) وسنناقش في هذا الفصل كل اختبار على حدة من حيث أهميته وطرق حسابه وأهم مجالات ومشاكل تطبيقه.

اختبار ستيودنت - «ت»[1]
(اختبار الفرق بين المتوسطات)

أوضحنا في الفصل السابق أن الاختبارات الاحصائية الخاصة بالعينات تفترض في أغلب الأحيان أن تكون بيانات العينة موزعة توزيعاً معتدلاً. وكما عرفنا أن توزيع العينة قد لا يكون كذلك، لذا فإن من الواجب إجراء بعض التعديلات في البيانات ليتسنى لها الاقتراب من التوزيع المعتدل. وذكرنا أيضاً أن بيانات العينة مهما كانت متماثلة فإنها لن تعكس تماماً خصائص المجتمع الذى سحبت منه، وبالتالي توجد بعض الفروق بين خصائص العينة ومعالم المجتمع الذى تمثله والتي يمكن تقديرها على أساس احتمالات أخطاء معينة (مستويات المعنوية α أو مستويات الثقة β).

وكما ذكرنا لا نستطيع دائماً قياس جميع المفردات في المجتمع لمعرفة معلمته (المتوسط الحسابي)، ولكن يستعاض عن متوسط المجتمع بمتوسط عينة حجمها كبير. وحتى يكون متوسط العينة ممثلاً ومماثلاً لمتوسط المجتمع يجب أن يكون الفرق بين المتوسطين صغيراً. أي إذا كان هناك فرض يقول أن متوسط العينة يساوي متوسط المجتمع فأننا نقبل هذا الفرض وذلك على العكس إذا كان الفرق بين المتوسطين كبيراً فإننا نرفض الفرض السابق حيث أنه في هذه الحالة سيكون هناك

---

(١) اكتشف العالم البريطاني William S. Gosset التوزيع الإحتمالي «ستيودنت - ت» في سنة ١٩٠٨ ولم يشأ أن يذكر إسمه فنشره بإمضاء ستيودنت (أى طالب Student) كبديل مستعار لإسمه، وأعطى الحرف الأخير في الكلمة وهو (ت = t) كأسم للاختبار الذى يستخدم هذا التوزيع في المعايرة الإحصائية.

اختلاف جوهرى بين متوسط العينة ومتوسط المجتمع. وتعنى كلمة الاختلاف الجوهرى Significant إحصائياً أن معلمة المجتمع تختلف اختلافاً كبيراً ولا تتفق مع نتائج العينة المسحوبة.

وبالمثل عند إجراء البحوث والدراسات الجغرافية تقابلنا كثيراً من المشاكل التى تتطلب المقارنة بين متوسطى عينتين لمعرفة ما إذا كانت هاتان العينتان مسحوبتين من مجتمعين مختلفين ولهما نفس المتوسط أم مسحوبتين من نفس المجتمع. فمثلاً قد نجد من الأفضل عملياً عند اختبار مدى فاعلية عامل معين أن نسحب عينتين الأولى لتمثل المجتمع قبل تأثير هذا العامل والثانية لتمثل المجتمع بعد تأثير العامل ونختبر ما إذا كان الفرق بين متوسط العينتين فرق جوهرى أو غير جوهرى، فإذا كان الفرق جوهرى نستنتج فاعلية العامل. أما إذا كان الفرق غير جوهرى فإننا نستنتج عدم فاعلية هذا العامل. وأن الفرق قد يكون راجعاً للصدفة المطلقة أو قد يكون ناتجاً من خطأ المعاينة. ولتوضيح ذلك نقول أنه عند استخدام نوعين مختلفين من المخصبات الزراعية لمعرفة ما إذا كان لهذين النوعين تأثيراً واضحاً على نوع معين من التربة فى منطقتين لهما نفس الظروف، وبالتالى على الإنتاج الزراعى أم لا، تسحب عينة من المجتمع الأول وعينة من المجتمع الثانى ثم يحسب المتوسط الحسابى لكل عينة على حدة، ثم يجرى الاختبار على قيم المتوسطين. فإذا كانت مَ١ ، مَ٢ هما متوسطى المجتمعين الأول والثانى على التوالى، وكانت سِ١ ، سِ٢ هما متوسطى العينتين المسحوبتين من المجتمعين السابقين فإن الفرق (سِ١ − سِ٢) هو متغير عشوائى الفرق (مَ١ − مَ٢). ويحدد الفرق بين المتوسطين بواسطة الوحدات المعيارية، أى بتحويل القيم الأصلية إلى وحدات معيارية ومعايرتها بوحدات معيارية بوحدات معيارية نظرية حتى نستطيع الحكم على أن هذا الفرق هو فرق جوهرى أم لا. وهناك مجالات أخرى مشابهة منها على سبيل المثال لا الحصر: اختبار وجود فروق جوهرية بين مستوى كفاءة عمال الإنتاج فى مصنعين مختلفين، أو اختبار وجود فروق جوهرية بين درجة صلابة نوعين من الصخور. فإذا ثبت أن الفرق بين كل عينتين هو فرق جوهرى أو حقيقى فإن ذلك يكون دليلا على أن العينتين مسحوبتان من مجتمعين مختلفين، أما إذا ثبت أن الفرق غير جوهرى فإن ذلك

يعنى أن الفرق بين متوسطى العينتين برجع الخطأ الصدفة أو لخطأ المعينة وأن العينتين قد تكونا مسحوبتان من مجتمع واحد أو من محتمعين لهما نفس المتوسط الحسابى.

ويستخدم أسلوب أو أختبار ستيودنت (ت) لاختبار المتوسطات فى حالة إذا لم يكن تباين المجتمع معلوماً والذى يستبدل بتباين العينات. وبما أن تباين أية عينة لايساوى بالضبط تباين المجتمع المسحوبة منه، لذلك فإننا لو استخدمنا توزيع «ز» فإن اختبار المتوسطات الحسابية سيتعرض للخطأ. غير أنه من الملاحظ على الاختبارين «ز»، «ت» أن قيم «ت» النظرية تصبح تقريباً نفس قيم توزيع «ز» إذا زاد حجم العينة عن ٩٠ مفردة. لذلك ففى الحالات التى يكون فيها حجم العينة أكبر من ٩٠ مفردة. لذلك ففى الحالات التى يكون فيها حجم العينة أكبر من ٩٠ مفردة فإنه يمكننا استخدام إما توزيع «ت» أو توزيع «ز» للاستدلال على صحة الفرض الموضوع لاختبار المتوسطات. كما يلاحظ أن قيمة «ت» تقاس أيضاً بوحدات الخطأ المعيارى للمجتمع المقدر من بيانات العينة ولذلك فإن توزيع «ت» يمثل توزيعات للمتوسطات الحسابية للعينات، ولكن نظر لأن توزيع قيم «ت» يعتمد على حجم العينة أو بالأحرى على درجات الحرية فانه يلزم عمل توزيع احتمالى لكل درجة من درجات الحرية والجدول المختصر فى ملاحق الكتاب الذى يبين قيمة «ت» عند مستويات معنوية مختلفة يعتبر كافياً لاستخدامه فى العينات الصغيرة. وتحسب درجات الحرية لعينة واحدة عد مفرداتها «ن» بطرح مفردة واحدة من هذه المفردات (أى ن – ١)، ويساعد ذلك فى تحديد تباين العينة والذى هو عبارة عن مجموع مربعات انحرافات القيم عن متوسطها الحسابى مقسوماً على درجات الحرية للعينة. ويمكن وضع ذلك فى الصيغة الرياضية الآتية:

$$\text{ع}^{2} = \frac{\text{مجـ}\,(\text{س} - \bar{\text{س}})^{2}}{\text{ن} - 1}$$

حيث ع^{2} هى تباين العينة، $(\text{س} - \bar{\text{س}})^{2}$ هى انحرافات القيم عن متوسطها الحسابى.

وهناك عدة شروط يجب توافرها عند استخدام إختبار «ت» وبغيرها فان النتائج التى نتوصل إليها لا تكون صحيحة:

أولاً: استقلال مفردات العينات قيد الاختبار أى أن كل عينة تكون مسحوبة بطريقة عشوائية ومستقلة عن الأخرى.

ثانياً: أن يكون التوزيع التكرارى للصفة المتغيرة لكل عينة توزيعاً معتدلاً.

ثالثاً: أن يكون هناك تجانس بين العينات، ويقصد هنا بالتجانس مدى التفاوت بين تباين أى عينتين، ويقاس هذا المدى بقسمة التباين الأكبر على التباين الأصغر، وليس بإيجاد الفرق بين تباين العينتين.

رابعاً: يجب لا يكون الفرق بين متوسطى العينتين كبيراً، وذلك لأن حجم العينة يؤثر على مستوى معنوية أو دلالة قيمة «ت»، وأن هذا المستوى يتأثر إلى حد كبير بدرجات الحرية الذى بدوره يحدد المنطقة الحرجة أو منطقة رفض فرض العدم الخاص باختبار.

اختبار المتوسطات للعينات الكبيرة (ن > ٣٠):

عند اختبار عينتين عشوائيتين (مستقلتين) من مجتمعين مختلفين وكان حجم العينة الأولى هو $ن_١$ ومتوسطها هو $\bar{س}$ وتباين مجتمعها هو $ع_١^٢$ وتباين المتوسط = $\frac{ع^٢}{ن_١}$ ، بينما كان حجم العينة الثانية هو $ن_٢$ وتباين مجتمعها هو $ع_٢^٢$ وتباين المتوسط = $\frac{ع_٢^٢}{ن_٢}$ ، فإن تباين الفرق بين المتوسطين يمكن حسابه من المعادلة الآتية:

$$ع^٢_{\bar{س}_١ - \bar{س}_٢} = \sqrt{\frac{ع_١^٢}{ن_١} + \frac{ع_٢^٢}{ن_٢}} \quad .... (٩ -١)$$

ويكون الانحراف المعيارى (الخطأ المعيارى) للفرق بين المتوسطين عبارة عن الجذر التربيعى للتباين أى أن:

$$ع_{س - س_٢} = \sqrt{\frac{ع_١^٢}{ن_١} + \frac{ع_٢^٢}{ن_٢}} \quad ....(٩ - ٢)$$

ويتم بذلك حساب مقياس (ت) من المعادلة الآتية:

$$ت = \frac{\text{الفرق بين متوسطى العينتين}}{\text{الخطأ المعيارى لهذا الفرق}}$$

$$ت = \frac{س_١ - س_٢}{\sqrt{\frac{ع_١^٢}{ن_١} + \frac{ع_٢^٢}{ن_٢}}} \quad ....(٩ - ٣)$$

وتتوزع القيمة التى نحصل عليها من هذا المقياس تبعاً (ت) بدرجات الحرية ($ن_١ + ن_٢ - ٢$)

وإذا كان حجم كل من العينتين $ن_١$، $ن_٢$ كبيراً وأن كلا منهما مسحوبة من مجتمع مختلف وحسبنا المتوسط الحسابى لكل عينة وكان الفرق بينهما هو ($س_١ - س$) = أ، وإذا سحبنا الفرق بين متوسطى العينتين ثم وضعنا هذه الفروق فى شكل توزيع تكرارى نجد أن متوسط هذه لفروق ($\frac{مجـ أ}{ن}$ حيث ن هى عدد العينات) يتوزع توزيعاً قريباً من التوزيع المعتدل. وإذا كان تباين المجتمعين متساويين ($ع_١^٢ = ع_٢^٢$) فإنه يمكن حساب التباين الكلى للمجتمع ثم الصيغة الإحصائية الآتية:

$$ع^٢ = \frac{مجـ (س_١ - س_١)^٢ + (س_٢ - س_٢)^٢}{ن_١ + ن_٢ - ٢}$$

أما إذا تساوت المفردات فى العينيتن (أى أن $ن_١ = ن_٢ = ن$) فإن الخطأ المعيارى للفرق بين المتوسطين سيكون فى هذه الحالة عبارة عن:

$$ع_{س_١ - س_٢} = \sqrt{\frac{٢ ع^٢}{ن}}$$

حيث $ع^{٢}$ هى التباين المشترك بين العينتين والذى يساوى:

$$ع^{٢} = \frac{\text{مجموع مربعات الإنحرافات للعينتين}}{\text{مجموع درجات الحرية}}$$

وبذلك يمكن حساب الخطأ المعيارى للفرق بين المتوسطين بالصيغة الاحصائية الآتية:

$$ع_{س^{٢}_{١} - س^{٢}_{٢}} = \sqrt{\frac{٢}{ن}\left[\frac{مجـ(س_{١} - \bar{س}_{١})^{٢} + مجـ(س_{٢} - \bar{س}_{٢})^{٢}}{ن^{٢}_{١} + ن^{٢}_{٢} - ٢}\right]}$$

وتكون قيمة (ت) المحسوبة فى هذه الحالة هى:

$$ت = \sqrt{\frac{٢}{ن}\left[\frac{مجـ(س_{١} - \bar{س}_{١})^{٢} + مجـ(س_{٢} - \bar{س}_{٢})^{٢}}{ن^{٢}_{١} + ن^{٢}_{٢} - ٢}\right]} \quad ..... (٩-٤)$$

وبمقارنة قيمة (ت) المحسوبة لكل من الحالتين والقيمة المناظرة لها فى جداول توزيع (ت) بمدلول درجات الحرية وعند مستوى معنوية معين يمكن رفض فرض العدم إذا كانت قيمة (ت) المحسوبة أكبر من قيمة (ت) النظرية. والعكس يكون صحيحاً (أى نقبل فرض العدم) إذا كانت قيمة (ت) المحسوبة أقل من قيمتها النظرية عند مستوى الدلالة (المعنوية) المعين.

مثال (١):

أخذت عينتين من إنتاج حقول الفحم فى منطقة ما لفترة عشرة سنوات، فكانت بيانتهما كمايلى:

	العينة الأولى	العينة الثانية
حجم العينة	١٠٠	١٠٠
الوسط الحسابى	٣٥	٢٤
الإنحراف المعيارى	٤	٥

والمطلوب اختيار الفرض القائل بأن متوسطى المجتمعين متساويات فى مقابل الفرض البديل القائل أن متوسط إنتاج حقول مجتمع العينة الأولى أكبر من متوسط مجتمع حقول العينة الثانية وذلك بمستوى معنوية ٠,٠٥ .

الفرض: H_0 : مَ١ = مَ٢ فى مقابل H_1 : مَ١ > مَ٢

أو H_0: (مَ١ − مَ٢) = صفر فى مقابل H_1: (مَ١ − مَ٢) ≠ صفر

وبحساب قيمة (ت) من العينة بافتراض أن مَ١ − مَ٢ = صفر نجد أن:

$$ت = \frac{٣٥ - ٣٤}{\sqrt{\frac{(٤)^٢}{١٠٠} + \frac{(٥)^٢}{١٠٠}}} = \frac{١}{\sqrt{\frac{١٦}{١٠٠} + \frac{٢٥}{١٠٠}}}$$

$$= \frac{١}{\sqrt{٠,٤١}} = \frac{١}{٠,٦٤} = ١,٥٦$$

وحيث أن قيمة (ت) المحسوبة من العينة (١,٥٦) أقل من القيمة (ت = ٢,١) المناظرة لدرجات الحرية (ن١ + ن٢ − ٢ = ١٠ + ١٠ − ٢ = ١٨) فى جدول توزيع (ت) بمستوى معنوية ٠,٠٥ فان القيمة المحسوبة لاتقع فى منطقة رفض الفرض، وعلى ذلك فان البيانات الخاصة بالعينتين قيد الاستقصاء غير كافية لرفض فرض العدم عند مستوى معنوية أو دلاله ٠,٠٥ ونستنتج من ذلك أن الفرق (س١ − س٢) هو فرق يرجع للصدفة المطلقة أو لأخطاء المعاينة وليس فرق جوهرياً بمستوى معنوية (α) = ٠,٠٥ .

اختبار المتوسطات للعينات الصغيرة (ن < ٣٠):

ذكرنا أنه فى حالة إذا كان الانحراف المعيارى للمجتمع غير معلوم وكانت العينة كبيرة الحجم (ن > ٣٠) فإنه يمكن حساب الخطأ المعيارى للمتوسط الحسابى للمجتمع من واقع الانحراف المعيارى المحسوب للعينة (عـ) بدلاً من

الانحراف المعيارى المجهول (ع) للمجتمع. ولكن قد نضطر إلى سحب عينة صغيرة (ن ‹ ٣٠) ومنها يمكن الحصول على تقدير غير متحيز للمعلمة (ع)، وذلك باستخراج الجذر التربيعى لخارج قسمة مجموع مربعات الانحرافات عن المتوسط الحسابى للعينة على عدد المتغيرات المستقلة الخطية (درجة الحرية ن −١)، لأن العينة الصغيرة تعطى معلومات عن المجتمع أقل دقة من بيانات العينات الكبيرة. رفى مثل هذه الحالة يتبع نفس الأسلوب الاحصائى السابق، أى يبنى الاستنتاج الاحصائى على أساس التوزيع الاحتمالى (ت)، فيصمم الاختبار الاحصائى لمقارنة الفرق بين متوسطى عينتين صغيرتين وذلك فى ضوء تقدير الخطأ المعيارى للفرق بين المتوسطين من واقع البيانات المشاهدة للعينتين المستقلتين بافتراض أنهما مسحوبتان مننفس المجتمع وفى هذه الحالة نستخدم الصيغة الآتية للقيمة المعيارية (ت):

$$\text{ت} = \frac{\bar{\text{س}}_{١} - \bar{\text{س}}_{٢}}{\sqrt{\frac{\text{ع}_{١}^{٢}}{\text{ن}_{١} - ١} + \frac{\text{ع}_{٢}^{٢}}{\text{ن}_{٢} - ١}}} \quad ....\ (٩ - ٥)$$

وهذه القيمة المعيارية لها توزيع احتمالى(ت) بدرجات الحرية ($\text{ن}_{١}$ + $\text{ن}_{٢}$ −٢) وبما أن قيمة الانحراف المعيارى للمجتمع (عـ) غير معلومة، فإنه يمكن الحصول على تقدير غير متحيز لهذه القيمة من واقع عينتين مستقلتين كمايلى:

$$\text{ع}_{\bar{\text{س}}_{١} - \bar{\text{س}}_{٢}} = \sqrt{\frac{\text{مجـ}(\text{س} - \bar{\text{س}})^{٢}}{\text{ن}_{١} - ١} + \frac{\text{مجـ}(\text{س} - \bar{\text{س}})^{٢}}{\text{ن}_{٢} - ١}} \quad .....\ (٩-٦)$$

$$\text{ولكن ع}^{٢} = \frac{\text{مجـ}(\text{س} - \bar{\text{س}})^{٢}}{\text{ن}}$$

$$\therefore \text{ن ع}^{٢} = \text{مجـ}(\text{س} - \bar{\text{س}})^{٢}$$

وبالتعريض عن مجـ $(\text{س} - \bar{\text{س}})^{٢}$ فى المعادلة (٩-٦) نجد أن:

$$\text{ع} = \sqrt{\frac{\text{ن}_{١}\,\text{ع}_{١}^{٢}}{\text{ن}_{١}} + \frac{\text{ن}_{٢}\,\text{ع}_{٢}^{٢}}{\text{ن}_{٢} - ٢}} \quad ....\ (٩ - ٧)$$

ولكن الخطأ المعيارى للعينة الأولى = $\frac{ع}{\sqrt{ن_١}}$ ، والخطأ المعيارى للعينة الثانية = $\frac{ع}{\sqrt{ن_٢}}$ ، والخطأ المعيارى للفرق بين متوسطى عينتين مسحوبتين من نفس المجتمع هو :

$$ع_{\bar{س}_١ - \bar{س}_٢} = \sqrt{\frac{ع^٢}{ن_١} + \frac{ع^٢}{ن_٢}} = ع \times \sqrt{\frac{١}{ن_١} + \frac{١}{ن_٢}} \quad ... (٩-٨)$$

وبالتعويض عن قيمة (ع) فى المعادلة (٩-٧) نجد أن:

$$ع_{\bar{س}_١ - \bar{س}_٢} = \sqrt{\frac{ن_١ ع_١^٢ + ن_٢ ع_٢^٢}{ن_١ + ن_٢ - ن}} \times \sqrt{\frac{١}{ن_١} + \frac{١}{ن_٢}} \quad ... (٩-٩)$$

وبذلك يمكن وضع الاختيار (ت) على الصورة المعيارية الآتية :

$$ت = \frac{\text{الفرق بين متوسطى العينتين}}{\text{الخطأ المعيارى للفرق}} = \frac{\bar{س}_١ - \bar{س}_٢}{\sqrt{\frac{ن_١ ع_١^٢ + ن_٢ ع_٢^٢}{ن_١ + ن_٢ - ن}} \times \sqrt{\frac{١}{ن_١} \quad \frac{١}{ن_٢}}}$$

(٩ - ١٠) ....

وفى حالة تساوى عدد المفردات لكل عينة من العينتين (أى أن ن١ = ن٢ = ن) فإن الصورة المعيارية للاختبار تكون على النحو التالى:

$$ت = \frac{\bar{س}_١ - \bar{س}_٢}{\sqrt{\frac{ع_١^٢ + ع_٢^٢}{ن - ١}}} \quad .... (٩ - ١١)$$

مثال (٢):

فى تجربة لمعرفة تأثير نوعين من السماد على محصول القطن سمدت ١٠ أفدنة من السماد الأول فكان متوسط إنتاجها ٧,٥ قنطاراً للفدان الواحد بانحراف

معيارى قدره ٠,٦ قنطاراً، كما سمدت ٩ أفدنة أخرى من نفس درجة خصوبة التربة بالسماد من النوع الثانى فانتجب فى المتوسط ٦,٨ قنطاراً للفدان الواحد بانحراف معيارى ٠,٥ قنطاراً، فهل يمكن الحكم بأن النوع الأول من السماد أفضل من النوع الثانى وذلك بمستوى معنوية ٠,٠٥ .

١- الفرض H_0 : $\bar{س}_١ \geq \bar{س}$

، الفرض H_1 : $\bar{س}_١ > \bar{س}$

٢- مستوى المعنوية = ٠,٠٥

٣- القيمة المعيارية للاختبار هى:

$$ت = \frac{\bar{س}_١ - \bar{س}_٢}{\sqrt{\frac{ن_١ ع_١^٢ + ن_٢ ع_٢^٢}{ن_١ + ن_٢ - ٢}} \times \sqrt{\frac{١}{ن_١} + \frac{١}{ن}}}$$

٤- توزيع ت الاحتمالى (النظرى) بدرجات الحرية (ن١ + ن٢ -٢) = (١٠ + ٩-٢ =١٧)

٥- منطقة الرفض: طبقاً للقواعد السابقة فانه يمكن رفض فرض العدم إذا كانت قيمة (ت) المحسوبة من البيانات المشاهدة أكبر من + ٢,١١ فى اختبار الطرفين، ولكن فى اختبار الطرف الواحد نرفض فرض العدم إذا كانت (ت) أقل من - ١,٧٤ .

٦- تحسب قيمة (ت) من البيانات المشاهدة للعينتين كمايلى:

$\bar{س}_١ = ٧,٥$ $\bar{س}_٢ = ٦,٨$

$ع_١ = ٠,٥$ $ع_٢ = ٠,٦$

$ن_١ = ١٠$ $ن_٢ = ٩$

$$ت = \frac{٧,٥ - ٦,٨}{\sqrt{\frac{١٠ (,٥)^٢ + ٩ (,٦)^٢}{١٠ + ٩ - ٢}} \times \sqrt{\frac{١}{١٠} + \frac{١}{٩}}}$$

$$= \frac{,٧}{,١٥٥} = ٤,٥١٦$$

٧- الاستنتاج: بما أن قيمة (ت) المحسوبة (٤,٥١٦) أكبر من قيمة (ت) النظرية (٢,١١) فمعنى ذلك أنها تقع ضمن منطقة الرفض (المنطقة الحرجة)، وبهذا نرفض فرض العدم ونقبل الفرض البديل أى أن هناك فرق جوهرى بين النوعين من السماد فى درجة تخصيبهما للتربة، ونستنتج أن النوع الأول من السماد ينتج لحصولا أوفر من النوع الثانى من السماد.

مثال (٣):

إذا كانت لدينا عينتين مستقلتين من عمال مصنعى أسمنت وكان عدد عمال العينة الأولى ٩ عاملاً وعدد عمال العينة الثانية ١٥ عاملاً وكان متوسط إنتاج عمال العينة الأولى ٥٠ طناً فى الشهر بانحراف معيارى ٧,٥ طن، ومتوسط إنتاج عمال العينة الثانية ٥٥ طناً فى الشهر وبانحراف معيارى ٦ طن، فهل يمكن الحكم على أن متوسط الإنتاج الشهرى فى المصنعين متساوى؟

١- الفرض H_0: $\bar{س}_١ = \bar{س}_٢$ ، الفرض H_1: $\bar{س}_١ \neq \bar{س}_٢$

٢- مستوى المعنوية = ٠,٠٥

٣- درجات الحرية للتوزيع الإحتمالى (ت) = ٩ + ١٥ - ٢ = ٢٥

٤- منطقة الرفض: تبعاً للفروض السابقة نجد أن قيمة (ت) النظرية التى تحدد منطقة الرفض هى ± ٢,٠٦ أى تقبل فرض العدم إذا كانت (ت) أقل من + ٢,٠٦ أو أكبر من -٢,٠٦.

٥- حساب قيمة (ت) من البيانات المشاهدة على النحو التالى:

$\bar{س}_١ = ٥٠$	$\bar{س}_٢ = ٥٥$
$ع_١ = ٧,٥$	$ع_٢ = ٦$
$ن_١ = ١٢$	$ن_٢ = ١٥$

$$\text{ت} = \frac{50 - 55}{\sqrt{\frac{12\,(7,5)^2 + 15\,(6)^2}{12 + 15 - 2} \times \sqrt{\frac{1}{12} + \frac{1}{15}}}}$$

$$= \frac{-5}{2,7} = -1,85$$

٦- الاستنتاج: بما أن قيمة (ت) المحسوبة (-١,٨٥) أكبر من قيمة (ت) النظرية (-٢,٠٦) فانها لا تقع فى منطقة الرفض (المنطقة الحرجة)، ولذلك نقبل فرض العدم القائل أن متوسط الإنتاج فى المصنعين متساوى، أى أنه لا يوجد فرق جوهرى بين المتوسطين وذلك بمستوى معنوية ٠,٠٥ وأن أى فرق بينهما هو فرق يرجع للصدفة المطلقة أو ينتج عن خطأ المعاينة.

ثانياً: تحليل التباين

Analysis of Variance

(اختبار - ف)

التباين هو أحد مقاييس التشتت التى عرفنا من قبل كيفية حسابها من قيم مفردة أو من جداول التوزيعات التكرارية. ولقد تأكدت أهمية التباين فى الدراسات والبحوث التى تقوم على أساس إحصائى كمى من حيث أنه المقياس الذى يوضح مدى التجانس أو الاختلاف لثلاث عينات أوأكثر ومدى صلتها بالمجتمع الاحصائى الذى تمثله. ولذا تعد طريقة تحليل التباين أشهر وأهم طرق التحليل الاحصائى للبيانات بعامة.

وفى الفصل السابق والقسم الأول من هذا الفصل تمت معرفة كيفية تحليل واختبار الفروق بين متوسطى عينتين للحكم على خصائص مجتمعيهما، أى أنه تحليل يتعلق بمجتمعين أو عينتين فقط. ولقد اعتمدنا فى تحديد مستويات المعنوية أو الدلالة الاحصائية لاختبار الفروق بين المتوسطات على قيم «ز» فى حالة معرفة تباين المجتمع وذلك لقبول أو رفض الفرض الموضوع، أو على قيم «ت» فى حالة عدم معرفة تباين لامجتمع أى على أساس معرفة تباين العينة. والسؤال الآن هو

كيف يتصرف الباحث إذا كان لديه أكثر من عينتين؟ وللإجابة على هذا التساؤل فان الباحث سيضطر للقيام بعمليات حسابية بين متوسطات العينات كل عينتين على حده. وتحدد العمليات الحسابية على أساس الصيغة التالية:

$$\frac{ن (ن - ١)}{٢}$$

حيث ن هى عدد العينات المطلوب حساب الفروق لمتوسطاتها.

وللتخلص من كثرة العمليات الحسابية يمكن المقارنة بين متوسطات أكثر من عينتين باستخدام طريقة أخرى وهى تحليل التباين للعينات[١] "Analysis of Variance"

تحليل التباين:

يعتمد تحليل التباين أساساً على حساب التباين بين العينات Variance between Samples والتباين داخل كل العينات مجتمعة Samples Variance. أما المقياس المستخدم للحكم على مستوى معنوية أو دلالة الفروق بين متوسطات العينات فهو ما يطلق عليه بقيمة ف F. وتقاس قيم ف النظرية من جداول خاصة موضوعة لهذا الغرض عن طريق تحديد درجات الحرية لكل تباسن على حدة بين العنيات وداخل العينات. ودرجات الحرية لتباين بين العينات عبارة عن هـ - ١ حيث «هـ» هى عدد العينات. أما درجات الحرية لتباين داخل العينات فهى (ى - هـ) حيث «ى» هى العدد الكلى للمفردات. فمثلاً إذا كان هناك ٦ عينات وكل عينة مكونة من ١٠ مفردات (قياسات)، فإن درجات الحرية فى هذه الحالة هى:

د ب درجات الحرية لتباين بين العينات = (هـ - ١) = (٦-١) = ٥

د د درجات الحرية لتباين داخل العينات = (ى - هـ) = (٦ ×١٠) -٦ = ٦٠ -٦ = ٥٤

---

(١) تسمى طريقة تحليل التباين فى بعض الأحيان باختبار نسبة ف F ratio test نسبة إلى عالم الاحصاء Fisher مكتشف هذه الطريقة.

وهناك عدة شروط أساسية يجب أن تتوافر عند استخدام طريقة تحليل التباين لعدة عينات وبغيرها تكون نتائج هذه الطريقة مضللة:

١- أن يكون توزيع مفردات أو قيم العينات متصفاً بصفة الاعتدالية أو أن انحرافها عن التوزيع المعتدل بسيطاً.

٢- أن يكون التباين لقيم المجموعات متجانساً أو متماثلاً، أى لا نوجد فروق بين تباين العينات المقارنة الا نتيجة للصدقة وذلك عن طريق مقارنة تباينات العينات.

٣- أن تكون العينات المطلوب تطبيق تحليل البتاين عليها ذات ظروف واحدة أو متجانسة.

وتجدر الإشارة هنا إلى أنه يجب قبل بدئ وضع العمليات الحسابية لتحليل التباين أن يوضع فرض لاختباره بهذا المقياس وهو فى هذه الحالة فرض العدم الذى ينص على أنه لا توجد فروق بين متوسطات وتباين العنيات الداخلة فى التحليل. ولاختيار هذا الفرض تحسب قيمة (ف) بين العنيات وداخل العينات باتباع الخطوات التالية:

١- حساب التباين بين العنيات عن طريق حساب المربعات بين العينات.

٢- حساب التباين داخل العينات عن طريق حساب المربعات داخل العينات.

٣- حساب نسبة (ف) عن طريق قسمة التباين الأكبر على التباين الأصغر.

٤- حساب درجات الحرية لاستخدامها فى الكشف عن مستوى الدلالة أو المعنوية الاحصائية لنسبة (ف) المحسوبة ومايقابلها من نسبة (ف) النظرية. فاذا كانت قيمة (ف) المحسوبة أقل من نظيرتها فى الجدول حسب مستوى المعنوية أو درجة الثقة المطلوبة فان الفرض يرفض بمعنى أن هناك فروق جوهرية بين متوسطات العينات. والجدول التالى يوضح صورة النتائج النهائية لهذا الاختبار.

جدول رقم (٩-١)

خطوات حساب تحليل التباين لأكثر من عينتين

مصدر التباين	مجموع المربعات	درجة		ف
١- بين العينات	مجـ مربع مجموع كل مجموعة / ن − ن م٢	عدد العينات −١	مجموع المربعات مقسوماً على درجات	التباين الأكبر مقسوماً على التباين
٢- داخل العينات	الفرق بين ١، ٣	عدد القيم عدد		
٣- المجموع الكلى	مجموع مربعات القيم − ن م٢	مجموع درجات.		

حيث أن (ن م٢) هى عامل التصحيح وهو يساوى أيضاً (مجـ س)٢ / ن حيث مجـ س٢ هو مجموع كل قيم المفردات للعينات و ن هى العدد الكلى لكل مفردات العينات.

التحليل الاحادى التصنيف للتباين:

فى هذا النوع من التحليل يمكن تقسيم الاختلافات الكلية إلى مصدرين هما: الاختلافات التى ترجع إلى قياسات العينات والأخرى ترجع إلى الأخطاء التجريبية Experimantal Error. وتتكون البيانات هنا من عدد من العينات المستقلة بكل منها مجموعة من القياسات (شكل رقم ٩-١).

	١	٢	٣	٤	٥
أ	•	•	•	•	•
ب	•	•	•	•	•
جـ	•	•	•	•	•
د	•	•	•	•	•

القياسات

شكل رقم (٩-١)

طريقة جمع البيانات للتحليل الاحادى للتباين

وكما هو واضح من الجدول رقم (٩-١) فانه لحساب قيمة (ف) نتبع الخطوات السابقة، والتى تعرف بطريقة التحليل الأحادى للتباين فى الاجابة على المثال الآتى:

فى مقارنة لمعرفة تباين الإنتاج الزراعى لمحصول الأرز فى خمس مناطق حسب إنتاج كل منها لمدة ٢٠ سنة كما هو موضح فى الجدول رقم (٩-٢). والمطلوب اختبار الفرض القائل بعدم وجود فروق بين متوسطات الإنتاج وذلك بمستوى معنوية ٠,٠٥.

جدول رقم (٩-٢)

الإنتاج السنوى لمحصول الأرز فى المناطق

أ، ب، حـ، د، هـ (طن/ فدان) لفترة ٢٠ سنة

أ	أ٢	ب	ب٢	جـ	جـ٢	د	د٢	هـ	هـ٢
٢,٥	٦,٢٥	٢,٠	٤,٠٠	٢,١	٤,٤١	٢,٧	٧,٢٩	٣,٤	١١,٥٦
٢,٤	٥,٧٦	٤,٩	٢٤,٠١	١,٥	٢,٢٥	٣,٠	٩,٠٠	٤,٤	١٩,٢٤
٢,٦	٦,٧٦	٣,٣	١٠,٨٩	١,٣	١,٦٩	٣,٢	٠,٢٤	٤,٠	١٦,٠٠
١,٧	٢,٨٩	٣,٨	١٤,٤٤	١,٢	١,٤٤	٢,٤	٥,٧٦	٤,٩	٢٤,٠١
٢,٠	٤,٠٠	٤,٠	١٦,٠٠	١,٦	٢,٥٦	٣,٤	١١,٥٩	٤,٩	٢٤,٠١
٢,٦	٦,٧٦	٣,٥	١٢,٢٥	٢,٢	٤,٨٤	٢,٨	٧,٨٤	٤,٠	١٦,٠٠
٢,٢	٤,٨٤	٢,٦	٦,٧٦	١,٨	٣,٢٤	٣,٧	١٣,٦٩	٤,٧	٢٢,٠٩
٥,١	٢٦,٠٠	٣,٥	١٢,٢٥	١,٧	٢,٨٩	٣,٥	١٢,٢٥	٤,٠	١٦,٠٠
٣,٢	١٠,٢٤	٣,٣	١٠,٨٨	٢,٧	٧,٢٩	٤,٧	٢٢,٠٩	٣,٤	١١,٥٦
٢,٢	٤,٨٤	٣,٨	١٤,٤٤	٣,٩	١٥,٢١	١,٩	٣,٦١	٢,٥	٦,٢٥
٣,٤	١١,٥٦	٣,٤	١١,٥٦	٢,٦	٦,٧٦	٢,٠	٤,٠٠	٧,٢	٥١,٨٠
٢,٠	٤,٠٠	١,٥	٢,٢٥	٢,٥	٦,٢٥	٢,٠	٤,٠٠	٤,٥	٢٠,٢٥
٢,٠	٤,٠٠	٢,٨	٧,٨٤	٢,١	٤,٤١	٢,٠	٤,٠٠	٣,٧	١٣,٦٩
٢,٣	٥,٢٩	٢,٠	٤,٠٠	٢,٢	٤,٨٤	٣,٦	١٢,٩٦	٥,٠	٢٥,٠٠
٣,٠	٩,٠٠	٢,٤	٥,٧٦	٢,٣	٥,٢٩	٢,٥	٦,٢٥	٦,٤	٤١,٠٠
٢,٥	٦,٢٥	٢,٧	٧,٢٩	١,٤	١,٩٦	١,٧	٢,٨٩	٤,٩	٢٤,٠١
١,٨	٣,٢٤	٢,٣	٥,٢٩	١,٣	١,٦٩	٢,٨	٧,٤٨	٣,٥	١٢,٢٥
٢,٤	٥,٧٦	٢,٩	٨,٤١	١,٤	١,٩٦	٢,٦	٦,٧٦	٩,٧	٩٤,٠٠
٢,٩	٨,٤١	٢,٠	٤,٠٠	١,١	١,٢١	٢,٦	٧,٧٦	٥,٠	٢٥,٠٠
٢,٢	٤,٨٤	٤,٠	١٦,٠٠	١,٥	٢,٢٥	٣,٠	٩,٠٠	٣,٢	١٠,٢٤
٥١,٠٠	٤٠,٦٥	٦٠,٧	١٩٨,٣٣	٣٨,٤	٨٢,٤٤	٥٦,١	١٦٧,٧٩	٩٣,٣	٤٨٤,١٧

الجموع (مجـ س)= ٥١,٠ + ٦٠,٧ + ٣٨,٤ + ٥٦,١ + ٩٣,٣= ٢٩٩,٥

$$\text{عامل التصحيح} = \frac{\text{(مجـ س)}^{٢}}{\text{ن}} = \frac{(٢٩٩,٥)^{٢}}{١٠٠} = ٨٩٧,٠٤$$

الجموع الكلي للمربعات = ١٤٠,٦٩ + ١٩٨,٣٣ + ٨٢,٤٤ + ١٦٧,٧٩ + ٤٨٤,١٧ − ٨٩٧,٠٤ = ١٠٧٣,٢٣ − ٨٩٧,٠٤ = ١٧٦,١٩

مجموع المربعات بين العينات = $\frac{١}{٢٠}$ [$(٥١,٠)^{٢}$ + $(٦٠,٧)^{٢}$ + $(٣٨,٤)^{٢}$ + $(٥٦,١)^{٢}$ + $(٩٣,٣)^{٢}$] − ٨٩٧,٠٤ = ٨٢,٩

مجموع المربعات داخل العينات =

الجموع الكلي للمربعات − مجموع المربعات بين العينات

= ١٧٦,١٩ − ٨٢,٩٠ = ٩٣,٢٩

درجات الحرية بين العينات = عدد العينات −١

= ٥ − ١ = ٤

درجات الحرية داخل عينات = عدد المفردات − عدد العينات

= ١٠٠ − ٥ = ٩٥

$$\text{التباين بين العينات} = \frac{\text{مجموع المربعات بين العينات}}{\text{درجات الحرية بين العينات}}$$

$$= \frac{٨٢,٩}{٤} = ٢٠,٧٣$$

$$\text{التباين داخل العينات} = \frac{\text{مجموع المربعات داخل العينات}}{\text{درجات الحرية داخل العينات}}$$

$$= \frac{٩٣,٢٩}{٩٥} = ٠,٩٨٢$$

نسبة ف = $\frac{\text{التباين الأكبر}}{\text{التباين الأصغر}}$ = $\frac{\text{التباين بين العينات}}{\text{التباين داخل العينات}}$

= $\frac{٢٠,٧٣}{٠,٩٨٢}$ = ٢١٠١١

وفى العادة توضح النتائج السابقة فى جدول كمايلى:

مصدر التباين	مجموع المربعات	درجات الحرية	التباين	ف
بين العينات	٨٢,٩٠	٤	٢٠,٧٣	٢١,١١
داخل العينات	٩٣,٢٩	٩٥	٠,٩٨٢	
المجموع	١٧٦,١٩	٩٩		

وبإستخدام جدول «ف» النظرية (F- Tables) وهو عبارة عن جدول لحساب نسبة التباين بدرجات بين العينات وداخل العينات وبمستوى معوية ٠,٠٥ وفى هذا الجدول تكون درجات الحرية الأفقية خاصة بدرجات الحرية للتباين الأصغر ودرجات الحرية الرأسية خاصة بدرجات الحرية للتباين الأكبر. وفى المثال بين أيدينا نجد أن نسبة «ف» لدرجات حرية (٤) بين العينات (ذات التباين الأكبر)، (٩٥) داخل العينات ذات التباين الأصغر عند مستوى دلالة ٠,٠٥ هى ٢,٤٩ تقريباً. وبما أن قيمة ف المحسوبة أكبر من القيمة النظرية فانه يمكن رفض فرض العدم الأساسى القائل بعدم وجود فروق جوهرية بين متوسطات إنتاج الأرز فى المناطق الخمسة، أو بمعنى آخر نقبل الفرض البديل وهو أنه توجد فروق جوهرية ذات دلالة إحصائية بين متوسطات الإنتاج. أى أن هناك احتمال مقداره ٠,٩٥ أن لا تكون الفروق بين المتوسطات قد حدثت بفعل الصدفة أو بطريقة عشوائية.

مثال آخر:

فى تجربة لدراسة المحتوى المائى (نسبة الرطوبة) لأربعة أنواع من التربة فى أحد المناطق أخذ من كل نوع عدد من العينات وتم استخلاص النتائج الموضحة فى

الجدول (٩-٣)، فهل يمكن الحكم بأن متوسطات المحتوى المائى (نسبة الرطوبة) للأربعة أنواع من التربة متساوية وذلك بمستوى دلاله ٠,٠٥ ، ٠,٠١ ؟

جدول رقم (٩-٣)

المحتوى المائى (نسبة الرطوبة) لأربعة أنواع من التربة فى منطقة ما

	أ	أ٢	ب	ب٢	جـ	جـ٢	د	د٢
	٣٣	١٠٨٩	٢٥	٦٢٥	١٧	٢٨٩	١٥	٢٢٥
	٢٦	٦٧٦	٢٨	٧٨٤	١٩	٣٦١	١٤	١٩٦
	٣٢	١٠٢٤	٢٥	٦٢٥	١٢	١٤٤	١٢	١٤٤
	٣٣	١٠٨٩	٢٤	٥٧٦	١٦	٢٥٦	١٣	١٦٩
			١٨	٣٢٤			١١	١٢١
المجموع	١٢٤	٣٨٧٨	١٢٠	٢٩٣٤	٦٤	١٠٥٠	٦٥	٨٥٥
عدد المفردات (ن)	٤		٥		٤		٥	
المتوسط	٣١		٢٤		١٦		١٣	

١- المجموع (س ن) = ٩٤ + ٩٠ + ٦٤ + ٦٥ = ٣٧٣

$$\text{٢- عامل التصحيح} = \frac{\text{مجـ (س ن)}^{٢}}{\text{ن}} = \frac{(٣٧٣)^{٢}}{١٨} = ٧٧٢٩,٣٩$$

٣- مجموع المربعات بين العينات =

$$\left[\frac{(١٢٤)^{٢}}{٤} + \frac{(١٢٠)^{٢}}{٥} + \frac{(٦٤)^{٢}}{٤} + \frac{(٦٥)^{٢}}{٥} \right] - ٧٧٢٩,٣٩$$

$$= ٨٥٩٣,٠٠ - ٧٧٢٩,٢٩ = ٩٨٧,٦١$$

٤- مجموع المربعات الكلية = [٣٨٧٨ + ٢٩٣٤ + ١٠٥٠ + ٨٥٥] -
٧٧٢٩,٣٩ = ٩٨٧,٦١

٥- مجموع المربعات داخل العينات = مجموع المربعات الكلية - مجموع المربعات بين العينات

= ٩٨٧,٦١ - ٨٦٣,٦١ = ١٢٤,٠٠

٦- درجات الحرية للمكونات السابقة:

درجات الحرية بين العينات = عدد العينات - ١ = ٤ - ١ = ٣

درجات الحرية الكلية = عدد المفردات - ١ = ١٨ - ١ = ١٧

درجات الحرية داخل العينات = عدد المفردات - عدد العينات

= ١٨ - ٤ = ١٤

(وهى تساوى أيضاً درجات الحرية الكلية - درجات الحرية بين العينات = ١٧ - ٣ = ١٤)

٧- التباين بين العينات = $\frac{٨٦٣,٦١}{٣}$ = ٢٨٧,٩

التباين داخل العنيات = $\frac{١٢٤,٠٠}{١٤}$ = ٨,٩

٨- قيمة (ف) = $\frac{\text{التباين الأكبر}}{\text{التباين الأصغر}}$ = $\frac{٢٨٧,٩}{٨,٩}$ = ٣٢,٥

وبتلخيص البيانات فى صورة جدولية نحصل على جدول تحليل التباين التالى:

جدول رقم (٩-٤)

تحليل التباين لمتوسطات المحتوى المائى (نسبة الرطوبة)

لأربعة أنواع من التربة

مصدر التباين	مجموع المربعات	درجات الحرية	التباين	قيمة (ف)
بين العينات	٨٦٣,٦١	٤ - ١ = ٣	٢٨٧,٩	٣٢,٥
داخل العينات	١٢٤,٠٠	١٨ - ٤ = ١٤	٨,٩	
المجموع	٩٨٧,٦١	١٨ - ١ = ١٧		

ولاختبار الفرض القائل بتساوى المتوسطات الأربعة نقارن قيمة (ف) المحسوبة بقيمة (ف) النظرية عند درجات حرية ٣، ١٤ ومنها يظهر من جداول ف – أن قيمة (ف) لمستوى دلالة ٠,٠٥ = ٢,٣٤ وقيمة (ف) لمستوى دلالة ٠,٠١ = ٥,٥٦. وحيث أن قيمة ف المحسوبة (٣٢,٥) أكبر من قيمة (ف) فى الجدول فإننا نرفض فرض العدم أى أن الاختلافات بين متوسطات العينات (أنواع التربة) هى اختلافات جوهرية حقيقية لا ترجع إلى الصدفة (أى أنها أكبر من الاختلافات العشوائية)، وذلك باحتمال ٠,٠٥ أو ٠,٠١ وبمعنى آخر أن متوسطات المحتوى المائى (نسبة الرطوبة) للأنواع الأربعة من التربة غير متساوية.

التحليل الثنائى التصنيف للتباين Two - way Anaiysis of Variance

إتضح من المثال السابق كيفية تحليل التباين لأكثر من مجموعتين بطريقة التحليل الأحادى التصنيف أو التحليل بالتصنيف فى اتجاه واحد. وهناك طريقة لتحليل التباين تعرف باسن التحليل الثنائى للتباين. وفى هذا النوع من التحليل يمكن تقسيم الاختلافات الكلية إلى ثلاث مصادر: إختلافات ترجع إلى العينات، اختلافات ترجع إلى المعاملات Treatements، واختلافات ترجع إلى الأخطاء

التجريبية Experimental . وتتكون البيانات هنا من عدد من العينات المستقلة بكل منها عدد من المفردات (المعاملات) كما فى الشكل رقم (٩-٢). ويوضح المثال التالى كيفية تطبيق طريقة الثنائى للتباين وخطوات العمل اللازمة معها:

١ ٢ ٣ ٤ ٥

المعاملات
١
٢
٣
٤
٥

شكل رقم (٩-٢)

طريقة أخذ البيانات لاختبارها بأسلوب التحليل الثنائى للتباين

مثال:

نفترض أننا سحبنا عينة عشوائية طبقية مكونة من ثلاث مجموعات من الأحواض الزراعية فى زمام أحد المراكز الإدارية بهدف مقارنة الإنتاجية الزراعية بين أنواع (طبقات) التربة التى تمثلها هذه المجموعات وهى التربة الطينية السوداء، والتربة الطينية الصفراء، والتربة الرملية. فإذا كانت كل مجموعة مكونة من عشرة أحواض تم إختيارها عشوائياً ثم رصدت لها كميات الإنتاج لمحاصيل الحبوب (بالطن) كما فى الجدول رقم (٩-٣)، فهل نقبل الفرض القائل بعدم وجود فروق بين متوسطات الإنتاج لهذه الأنواع الثلاثة من التربة.

جدول رقم (١٢-٣)
الإنتاجية الزراعية لمحاصيل الحبوب (بالطن)
لمجموعة من الأحواض الزراعية فى ثلاثة أنواع من التربة

	التربة الطينية		التربة الطينية		التربة الرملية		المجموع	
	أ	أ^{٢}	ب	ب^{٢}	جـ	ج^{٢}	(س ن)	ـس ن)^{٢}
	٢٤	٢٤	٢٤	٢٤	٢٤	٢٤	٢٤	٢٤
	٢٧	٢٧	٢٧	٢٧	٢٧	٢٧	٢٧	٢٧
	٢١	٢١	٢١	٢١	٢١	٢١	٢١	٢١
	٢٢	٢٢	٢٢	٢٢	٢٢	٢٢	٢٢	٢٢
	٢٦	٢٦	٢٦	٢٦	٢٦	٢٦	٢٦	٢٦
	١٩	١٩	١٩	١٩	١٩	١٩	١٩	١٩
	٢٥	٢٥	٢٥	٢٥	٢٥	٢٥	٢٥	٢٥
	٢٩	٢٩	٢٩	٢٩	٢٩	٢٩	٢٩	٢٩
	٢٦	٢٦	٢٦	٢٦	٢٦	٢٦	٢٦	٢٦
	٢٤	٢٤	٢٤	٢٤	٢٤	٢٤	٢٤	٢٤
المجموع	٢٤٣	٢٤٣	٢٤٣	٢٤٣	٢٤٣	٢٤٣	٢٤٣	٢٤٣

١- المجموع (مجـ $س_{ن}$) = ٢٤٣ + ٢٢٢ + ٢١٠ = ٦٧٥

٢- عامل التصحيح = $\frac{(مجـ\ س_{ن})^{٢}}{ن}$ = $\frac{(٦٧٥)^{٢}}{٣٠}$ = ١٥١٨٧,٥

٣- مجموع مربعات بين العينات (المجموعات) = $\frac{٤٥٨٥٣}{٣}$ − ١٥١٨٧,٥

= ٩٦,٨٣

٤- مجموع المربعات بين المعاملات

= $\frac{1}{10}$ [(٢٤٣)^{2} + (٢٢٢)^{2} + (٢١٠)^{2}] − ١٥١٨٧,٥

= ١٥٢٤٣,٣ − ١٥١٨٧,٥ = ٥٥,٨

٥- المجموع الكلي للمربعات = ١٥٤٨٩,٠ − ١٥١٨٧,٥ = ٣٠١,٥

٦- نحسب الخطأ التجريبي Experimental Error لمجموع المربعات (الفرق أو الفائض Resibual) وهو عبارة عن الفرق بين المجموع الكلي للمربعات ومجموع المربعات بين العينات ومجموع المربعات بين المعاملات كمايلي:

خطأ مجموع المربعات = ٣٠١,٥ − ٩٦,٨٣ − ٥٥,٨ = ١٤٨,٨٧

٧- نحسب درجات الحرية للتحليل كمايلي:

درجات الحرية بين المعاملات = عدد المعاملات (بين الصفوف) −١

= ١٠ − ١ = ٩

درجات الحرية بين العينات (بين الأعمدة) = عدد العينات −١

= ٣ − ١ = ٢

درجات الحرية للمفردات الكلية (ن) = عدد المفردات «ن» −١

= ٣٠ − ١ = ٢٩

درجات الحرية للخطأ (الفرق) = درجات الحرية للمفردات − درجات حرية التباين بين العينات − درجات الحرية بين المعاملات (بين الصفوف)

= ٢٩ − ٢ − ٩ = ١٨

أو

= (درجة الحرية بين العينات) × (درجات الحرية بين المعاملات)

= ٢ × ٩ = ١٨

ويمكن تلخيص النتائج السابقة في الجدول التالي:

جدول رقم (٩-٤)

جدول التحليل الثنائى للتباين 2 WAY - ANOVA

مصدر التباين	مجموع المربعات	درجات الحرية	التباين	قيمة ف
بين العينات	٩٦,٨٣	٢	٤٨,٤١٥	
بين المعاملات	٥٥,٨٠	٩		٥,٨٥
الخطأ (داخل المعاملات)	١٤٨,٨٧ .	١٨	٨,٢٧	
المجموع الكلى	٣٠١,٥٠	٢٩		

وبالبحث فى جداول (ف) لإيجاد قيمتها النظرية عند درجات الحرية ٢ بين المجموعات و ١٨ للخطأ عند مستوى دلالة ٠,٠٥ نجد أن = ٣,٥٥. وحيث أن قيمة (ف) المحسوبة فى المثال (٥,٨٥) أكبر من القيمة النظرية عند مستوى الدلالة ٠,٠٥ فإننا يجب أن نرفض الفرض القائل أنه لا توجد فروق جوهرية بين متوسطات المجموعات الثلاث من التربة من حيث الإنتاجية الزراعية للحبوب. وهذا يعنى أن هناك احتمالاً قدرة ٠,٩٥ أن الفروق بين الإنتاجية الزراعية للحبوب لهذه الأنواع من التربة لم تحدث بفعل الصدفة المطلقة أو أنها نتيجة خطأ عشوائى فى المعاينة ولكنها فروق حقيقية لها دلالة إحصائية.

الفصل العاشر
أساليب المقارنة اللابارامترية
(اللامعلمية)



أساليب المقارنة اللابارامترية (اللامعلمية)

عرضنا فى الفصل السابق الأساليب الكمية البارامترية التى تحلل المعنوية الاحصائية للاختلافات أو الفروق بين بيانات التوزيعات العينية والتوزيع المعتدل للمجتمع، وكان الاهتمام منصباً على مقارنة قيم المتوسط الحسابى والتباين المحسوبة مع عينتين أو أكثر لمتغير واحد. وذكرنا أيضاً أنه فى بعض الحالات لاتكون خصائص أو معالم المجتمع معلومة، وبالتالى لابد أن تستخدم أساليباً أخرى لإجراء إختبارات المعنوية وتحليل البيانات الأصلية التى لاتتصف باعتدالية توزيعاتها التكرارية حتى بعد تطبيق إحدى الطرق السابق ذكرها لتحويلها إلى توزيعات معتدلة.

وستتناول فى هذا الفصل دراسة الأساليب الكمية اللابارامترية (اللامعلمية) التى تستخدم فى مجال المقارنة بين مجموعتين أو أكثر من البيانات لمتغير واحد، وهى: مربع كاى، اختيار كولموجوروف – سميرنوف (اختبار «د» اختبار مان – هويتنى (اختبار «ى»)، اختبار ويلكوكسون، اختبار كروسكال – واليس (اختبار «هـ»). وتجدر الإشارة إلى أن هذه الاختبارات لاتشترط أن تكون البيانات من بيانات الفترة فقط ولكن يمكن تطبيقها أيضاً على البيانات الإسمية (التصنيفية أو النوعية) والترتيبية سواء كانت فى صورة قياسات فردية أو ثنائية ينتج عنها قيم عددية أو تكرارات أو رتب.

أولاً: اختبار مربع كاى (x2)

Chi - Square Test

يستخدم مربع كاى أساساً فى قياس مدى التطابق بين توزيعين أحدهما توزيع فعلى لمتغير تم قياسه والآخر توزيع نظرى أو متوقع. وعلى ذلك فوجه المقارنة يكون بين مجموعتين من البيانات التكرارية أحداهما فعلية والأخرى نظرية. ويكون الفرض الموضوع هو المتعلق بالفروق أو الاختلافات بين التوزيعات الفعلية أو المشاهدة والتوزيعات المتوقعة للوقوف على معرفة نوع هذه الفروق: هل هى فروق معنوية أو جوهرية، أم أنها مجرد فروق ظاهرية؟ فإذا كانت الفروق حقيقية فإن ذلك يعنى أنها نتيجة لعوامل مسئولة عنها وليست مرتبطة بعوامل أخرى مسببة لها، أما إذا كانت الفروق غير جوهرية فإن ذلك يعنى أنها نتيجة للصدفة المطلقة أو أنها نتيجة لمجموع العوامل غير المتحكم فيها أو ما يطلق عليه بالاختلافات والأخطاء العشوائية Random Errors. وبصفة عامة فإنه يمكن القول أن تحليل البيانات بواسطة مربع كاى يتم لتحقيق هدفين رئيسيين هما:

١- تحديد دلالة العلاقة بين مجموعتين أو أكثر من التصنيفات بالنسبة إلى خصائص العينة.

٢- تحديد دلالة انحرافات التكرارات الفعلية (المشاهدة) عن التكرارات المتوقعة، أو بمعنى آخر الحكم على مدى ملائمة النموذج المتوقع لتوزيع التكرارات الفعلية عن طريق اتباع اختبار جودة التوفيق Goodness of fit Test.

وكما سبق أن شاهدنا أنه فى عديد من المرات لاتتفق النتائج التى نحصل عليها من العينات فى جميع الحالات مع النتائج المتوقعة طبقاً لقواعد الاحتمالات. فلو افترض أنه فى عينة معينة لوحظ أن مجموعة من الحادثات الممكنة: هـ١، هـ٢، هـ٣ ... هـن تحدث بتكرارات ش١، ش٢، ش٣، ... ش ن التى تسمى بالتكرارات المشاهدة، وأنه طبقاً لقواعد الاحتمالات فإنه يتوقع أن تحدث بتكرارات ق١، ق٢، ق٣ ... قن والتى تسمى بالتكرارات المتوقعة كما فى الشكل التالى:

الحادث	هـ١	هـ٢	هـ٣ .... هـن
التكرار المشاهد	ش١	ش٢	ش٣ .... شن
التكرار المتوقع	ق١	ق٢	ق٣ .... قن

ويتطلب إجراء اختبار «مربع كاى» توفر الشروط الأساسية الآتية:

أولاً: أن لايقل العدد الكلى للقيم (حجم العينة) عن ٢٠.

ثانياً: أن تكون المشاهدات مستقلة عن بعضها البعض (أى لايؤثر اختيار أحد المفردات على اختيار المفردات الأخرى).

ثالثاً: أن تكون البيانات المشاهدة فى شكل قيم عددية أو تكرارات قائمة على العد فى كل فئة من الفئات، ولاتكون هذه البيانات فى صورة نسب مئوية على الإطلاق.

رابعاً: إذا كانت العينة تنقسم إلى فئتين فقط فيجب أن لايقل عدد التكرارات المتوقعة لهما عن ٥ تكرارات. إذا زاد عدد الفئات عن فئتين فيجب أن لايقل ١/٥ التكرارات المتوقعة، على الأكثر، عن ٥ تكرارات، ولكن يجب أن لايقل أبداً عدد التكرارات المتوقعة لأية فئة عن تكرار واحد. فإذا لم يتحقق ذلك فى العينات المقارنة فإنه يمكن ادماج فئتين أو أكثر فى فئة واحدة حتى يمكن أن تتم إجراءات المقارنة بأداة مربع كاى.

ولحساب قيمة مربع كاى تستخدم الصيغة الاحصائية الآتية:

$$\text{مربع كاى } (x^2) = \frac{(\text{ش}_1 + \text{ق})^2}{\text{ق}} + \dots\dots + \frac{(\text{ش}_\text{ن} + \text{ق})^2}{\text{ق}}$$

حيث (ش) هى التكرارات المشاهدة (ق) هى التكرارات المتوقعة أو النظرية. وتتراوح قيمة مربع كاى من صفر إلى α ويتوقف توزيعها على درجات الحرية. فإذا كانت النتيجة النهائية (القيمة المحسوبة) لاختبار مربع كاى أقل من نظيرتها فى توزيع مربع كاى[1] فى الجداول الاحصائية الخاصة به فى مستوى دلالة معين (α) فإننا نقبل فرض العدم أو بمعنى آخر أنه لايوجد اختلافات أو فروق معنوية بين التوزيعين المشاهد (الفعلى) والمتوقع، أما إذا كانت قيمة مربع كاى المحسوبة أكبر من مثيلتها النظرية فى جدول توزيع مربع كاى فإننا نرفض فرض العدم، وهذا يعنى وجود فروق معنوية أو حقيقية بين التوزيعين أى أن هناك عوامل غير عامل الصدفة لها تأثير على هذه الفروق. وتقوم المقارنة بين قيمة مربع كاى المحسوبة والنظرية، على أساس درجات الحرية (ن – ١) حيث ن هى عدد المجموعات أو الفئات. هذا فى حالة التصنيفات الأحادية أما فى حالة التصنيفات الثنائية كجداول الاقتران Contingency Tables فتحسب درجات الحرية كما يلى:

(د– ١) (– ١) حيث د = عدد الأعمدة، ص = عدد الصفوف

التحليل الأحادى The one - Sample Case

يتصل هذا التصنيف بالبيانات المشاهدة التى يتم تقسيمها حسب صفة واحدة من الصفات.

مثال:

جمعت عينة عشوائية لعدد ٦٠٠ قطعة صخرية من منطقة شاطئية، ووجد أن ١٨٠ قطعة منها من نوع الحجر الجيرى، ١٨٦ من الجرانيت والباقى ٢٣٤ من نوع الحجر الرملى. وذلك على الرغم من أن الأنواع الصخرية الثلاثة متمثلة فى المنطقة بمساحات متساوية، والفرض المراد اختباره فى هذه الحالة: حيث أن الأنواع

---

(١) توزيع مربع كاى النظرى غير منتظم، متوسطه يساوى درجات الحرية وتباينة يساوى ضعف درجات الحرية، ويقرب من التوزيع المعتدل كلما زادت درجات الحرية.

الرئيسية للصخور مساحتها متساوية فى المنطقة فإننا نتوقع أن يكون لكل نوع منها عدداً من القطع مماثلاً للنوع الآخر. والفرض البديل اختباره فى هذه الحالة: حيث أن الأنواع الرئيسية للصخور مساحتها متساوية فى المنطقة فإننا نتوقع أن يكون لكل نوع منها عدداً من القطع مماثلاً للنوع الآخر. والفرض البديل لذلك هو وجود فروق جوهرية فى عدد القطع الصخرية لكل نوع من أنواع الصخور. وحيث أن أعداد القطع قد جمع فى عينة واحدة فإنه ليس من الصواب أن نتوقع تساوى أعداد القطع لكل نوع، ولكن بما أن العينة عشوائية فإن أعداد القطع لاتبعد كثيراً عن التساوى من بعضها البعض، وفى هذه الحالة فإن مربع كاى يستخدم لتقدير الاحتمال القائل أن العينة قد أخذت من مجتمع تتساوى فيه أعداد القطع لكل نوع صخرى مع العلم بأن مستوى الدلالة المطلوب هو ٠,٠١ ويمكن ترتيب البيانات السابقة كما هى الحال فى جدول المقارنة التالى:

جدول رقم (١٠ – ١)

حساب مربع كاى بطريقة التحليل الأحادى

الحادث	عدد القطع المجموعة (التكرارات المشاهدة) ش	عدد القطع المتوقعة (ق)
حجر جيرى	١٨٠	٢٠٠
جرانيت	١٨٦	٢٠٠
حجر رملى	٢٣٤	٢٠٠

$$\text{مربع كاى} = \frac{(\text{ش}_1 + \text{ق})^2}{\text{ق}} + \frac{(\text{ش}_\text{ن} + \text{ق})^2}{\text{ق}}$$

$$= \frac{(180 - 200)^2}{200} + \frac{(186 - 200)^2}{200} + \frac{(234 - 200)^2}{200}$$

= ٢ + ,٩٨ + ٥,٧٨

= ٨,٧٦

درجات الحرية = عدد الفئات − ١ = ٣ − ١ = ٢

وبما أن قيمة مربع كاى من التوزيع النظرى المدون فى الجداول الخاصة به تساوى ٩,٧١ لدرجة حرية ٢ ولمستوى دلالة ٠١, فمعنى ذلك أنه لابد من قبول فرض العدم القائل: أنه يتوقع جمع عدد متساوى للقطع الصخرية من كل نوع من أنواع الصخور فى المنطقة بمستوى دلالة ١٪، وبناء عليه فإن الفرق بين أعداد القطع فى العينة هو فرق يرجع إلى الصدفة المحضة أو نتيجة خطأ المعاينة.

التحليل (التصنيف) الثنائى The Two Sample Case

لاحظنا فى التحليل الأحادى أن الاختبار للفرض الموضوع كان يتعلق بالفروق بين التكرارات المشاهدة والتكرارات المتوقعة. أما فى حالة التحليل الثنائى بمربع كاى فإن الفرض يتعلق باختبار الاستقلال بين صفتين أو تصنيفين أو اختبار عدم وجود علاقة بين هاتين الصفتين أو التصنيفين. ويحسب مربع كاى فى حالة وجود عينتين وفئتين على النحو التالى:

$$\frac{\text{ن}\left[(\text{أ د} - \text{ب حـ}) - \frac{\text{ن}}{٢}\right]^{٢}}{(\text{أ} + \text{ب})\,(\text{ح} + \text{د})\,(\text{أ} + \text{جـ})\,(\text{ب} + \text{د})}$$

حيث **ن** = العدد الكلى لمفردات العينتين (المجموع الكلى للتكرارات)، أ، ب، جـ، د = التكرارات لكل فئة من فئتى العينتين داخل «جدول الاقتران» كما هى الحال فى الجدول التالى:

أ	ب	(أ + ب)
جـ	د	(جـ + د)
(أ + د)	(ب + د)	

وتحسب درجات الحرية لهذا الاختبار من عدد الأعمدة وعدد الصفوف في الجدول كما يلي: .

(د - ١) × (ص - ١) حيث د = عدد الأعمدة، ص = عدد الصفوف

مثال (٥):

أخذت عينة لأسبوع واحد من سجلات إدارة مرور لمن تقدموا لامتحان قيادة السيارات فوجد أنه من بين ١٠٠ من النساء اجتاز منهن الامتحان لأول مرة ٧٥ ورسب ٢٥ . وأن من بين ١٥٠ رجلا اجتاز منهم الامتحان لأول مرة ٦٠ ورسب ٩٠ ، فإذا أخذت العينة السابقة على أساس أنها تمثل قطاعاً عرضياً لمجتمع المتقدمين لامتحان قيادة السيارات، فإن عدد النساء والرجال، كل على حدة، يمكن أن يعامل على أنه عينة عشوائية يهدف تحليلها بمربع كاى إلى معرفة ما إذا كانت هناك اختلافات جوهرية بين مقدرة كل من النساء والرجال على القيادة واجتيازهم اختبار القيادة في أول محاولة بمستوى ثقة ٩٥ ٪. وجدول الاقتران التالى يوضح بيانات هذا المثال.

جدول رقم (١٠ - ٢)
أعداد المتقدمين لامتحان قيادات السيارات
حسب النوع ونتيجة الامتحان

نتيجة الامتحان / النوع	ناجح	راسب	المجموع
نساء	٧٥ (أ)	٢٥ (جـ)	١٠٠
رجال	٦٠ (ب)	٩٠ (د)	١٥٠
المجموع	١٣٥	١١٥	٢٥٠

ومن جدول الاقتران السابق يمكن ايجاد قيمة مربع كاى كما يلي:

$$\text{مربع كاى} = \frac{\text{ن} \left[(\text{أ د} - \text{ب حـ}) - \frac{\text{ن}}{٢} \right]^{٢}}{(\text{أ} + \text{ب})\ (\text{جـ} + \text{د})\ (\text{أ} + \text{جـ})\ (\text{ب} + \text{د})}$$

$$= \frac{٢٥٠ \left[(٧٥ \times ٩٠ - ٦٠ \times ٢٥) - \frac{٢٥٠}{٢} \right]^{٢}}{(٧٥ + ٦٠)\ (٢٥ + ٩٠)\ (٧٥ + ٢٥)\ (٦٠ + ٩٠)}$$

$$= \frac{٢٥٠ \left[(٦٧٥٠ - ١٥٠٠) - ١٢٥ \right]^{٢}}{١٣٥ \times ١١٥ \times ١٠٠ \times ١٥٠}$$

$$= \frac{٢٥٠ \left[٥١٢٥ \right]^{٢}}{٢٣٢٨٧٥٠٠٠} = ٢٨,١٩٦$$

وحيث أن درجات الحرية = (د − ١) × (ص − ١)

= (٢ − ١) × (٢ − ١) = ١

فإن قيمة مربع كاى من توزيعها فى الجدول بدرجة حرية ١ ومستوى ثقة ,٩٥ (مستوى دلالة ٠,٠٥) هى ٠٣,٨٤ وبما أن قيمة مربع كاى المحسوبة هى ٢٨,١٩٦ أكبر من القيمة النظرية لمربع كاى فإننا نرفض فرض العدم. ومعنى ذلك أن الاختلاف فى مقدرة كل من النساء والرجال فى اجتياز امتحان القيادة فى أول محاولة لايرجع اطلاقاً إلى الصدفة فى العينة، بل أن بيانات العينة تعكس بكل دقة الاختلاف الحقيقى والجوهرى بين مقدرة النساء والرجال فى مجتمع السائقين.

ومما تجدر الإشارة إليه هنا أن معادلة «مربع كاى» المستخدمة لتحليل عينة واحدة يمكن تطبيقها أيضاً فى تحليل بيانات عينتين أو أكثر يمكن تقسيمها إلى أكثر من فئتين.

مثال (٢):

أجريت دراسة على عينة عشوائية من ١٥٠ سفحاً «منحدراً» طبيعياً أخذت من

ثلاث مناطق متجاورة تتصف كل منها بتركيب صخرى متميز (مارل فى المنطقة الأولى، حجر جيرى فى المنطقة الثانية، حجر رملى فى المنطقة الثالثة) لتحديد ما إذا كان هناك اختلاف جوهرى بين درجات انحدار السفوح ونوع التركيب الصخرى الذى تنشأ فوقه هذه السفوح. وقد وضع فرض العدم فى هذه الحالة وهو أنه لايوجد اختلاف جوهرى بين درجات انحدار السفوح الطبيعية التى تنشأ فوق أنواع مختلفة من التركيب الصخرى، وذلك فى مقابل الفرض البديل القائل بأنه على افتراض أن التاريخ الجيولوجى والغطاء النباتى ومقدار الاتجاه للسفوح متشابهة فى المناطق الثلاث فإنه يمكن أن نستنتج أن الاختلافات الجوهرية فى درجات الانحدار ما هى إلا نتيجة واختلاف طبيعة التركيب الصخرى نفسه، وأن أى اختلاف آخر إنما يرجع إلى الصدفة وذلك فى مستوى دلالة ٠,٠٥ ولاختبار الفرض السابق فإن أعداد السفوح لكل نوع من أنواع التركيب الصخرى «أى فى المناطق الثلاث» فى العينة قد قسمت إلى خمس فئات حسب درجة الانحدار، ووضعت البيانات الخاصة بذلك فى جدول الإقتران التالى:

جدول رقم (١٠ – ٣): التوزيع التكرارى لأعداد السفوح الطبيعية فوق ثلاثة أنواع مختلفة من التركيب الصخرى

الفئـــات	التكرارات (عدد السفوح)			المجموع
(درجة الانحدار)	مارل	حجر جيرى	حجر رملى	
صفر –	٢١	٩	٥	٣٥
٥ –	١٨	١١	٧	٣٦
١٠ –	٩	١٥	٩	٣٣
١٥ –	٨	١٣	١١	٣٢
٢٠ إلى أقل من ٢٥	١	٥	٨	١٤
المجموع	٥٧	٥٣	٤٠	١٥٠

فإذا كانت توقعاتنا بأن نسبة عدد السفوح فى كل فئة من فئات درجة الانحدار يتناسب مع العدد الكلى السفوح فمعنى ذلك أنه من المفروض أن من مجموع ٥٧ سفحاً طبيعياً تكونت فوق صخور مارلية ٢٣,٢ ٪ منها (أى ١٣,٣ سفحاً) ذات انحدارات تتراوح درجتها بين الصفر وأقل من ٥ درجات، ٢٤ ٪ منها (أى ١٣,٧ سفحاً) ذات انحدارات تتراوح درجتها بين أكثر من ٥ درجات وأقل من ١٠ درجات ٢٢ ٪ منها (أى ١٢,٢ سفحاً) ذات انحدارات تتراوح درجتها بين أكثر من ١٥ درجة وأقل من ٢٠ درجة، وأخيراً ٩,٣ ٪ منها (أى ٥,٣ سفحاً) ذات انحدارات تتراوح درجتها بين أكثر من ٢٠ درجة وأقل من ٢٥ درجة. وبالمثل يمكن ايجاد العدد المتوقع للتكرارات داخل كل فئة من الفئات للأنواع الثلاثة من التركيب الصخرى كما فى الجدول التالى:

جدول رقم (١٠ – ٤): حساب التكرارات المتوقعة من البيانات المشاهدة لعدد ١٥٠ سفحاً طبيعياً

الفئات (درجة الانحدار)	التكرارات (عدد السفوح)			المجموع
	مارل	حجر جيرى	حجر رملى	
صفر –	$\frac{٣٥ \times ٥٧}{١٥٠}$ (١٣,٣)	$\frac{٣٥ \times ٥٣}{١٥٠}$ (١٢,٤)	$\frac{٣٥ \times ٤٠}{١٥٠}$ (٩,٣)	٣٥
٥ –	$\frac{٣٦ \times ٥٧}{١٥٠}$ (١٣,٧)	$\frac{٣٦ \times ٥٣}{١٥٠}$ (١٢,٧)	$\frac{٣٦ \times ٤٠}{١٥٠}$ (٩,٦)	٣٦
١٠ –	$\frac{٣٣ \times ٥٧}{١٥٠}$ (١٢,٥)	$\frac{٣٣ \times ٥٣}{١٥٠}$ (١١,٧)	$\frac{٣٣ \times ٤٠}{١٥٠}$ (٨,٨)	٣٣

١٥ -	$\frac{٣٢ \times ٥٧}{١٥٠}$ (١٢,٢)	$\frac{٣٢ \times ٥٣}{١٥٠}$ (١١,٣)	$\frac{٣٢ \times ٤٠}{١٥٠}$ (٨,٥)	٣٢
٢٠ إلى أقل من ٢٥	$\frac{١٤ \times ٥٧}{١٥٠}$ (٥,٣)	$\frac{١٤ \times ٥٣}{١٥٠}$ (٤,٩)	$\frac{١٤ \times ٤٠}{١٥٠}$ (٣,٨)	١٤
المجمـــوع	٥٧	٥٣	٤٠	١٥٠

ولاختبار فرض العدم فى هذا المثال تطبق الصيغة التالية:

$$\text{مربع كاى} = \text{مجـ} \frac{(\text{ش ن} - \text{ق})^{٢}}{\text{ق}}$$

$$= \frac{(٢١ - ١٣,٣)^{٢}}{١٣,٣} + \frac{(٩ - ١٢,٤)^{٢}}{١٢,٤} + \frac{(٥ - ٩,٣)^{٢}}{٩,٣}$$

$$+ \frac{(١٨ - ١٣,٧)^{٢}}{١٣,٧} + \frac{(١١ - ١٢,٧)^{٢}}{١٢,٧} + \frac{(٩ - ١٢,٥)^{٢}}{١٢,٥}$$

$$+ \frac{(١٥ - ١١,٧)^{٢}}{١١,٧} + \frac{(٩ - ٨,٨)^{٢}}{٨,٨} + \frac{(٨ - ١٢,٢)^{٢}}{١٢,٢}$$

$$+ \frac{(١٣ - ١١,٣)^{٢}}{١١,٣} + \frac{(١١ - ٨,٥)^{٢}}{٨,٥} + \frac{(١ - ٥,٣)^{٢}}{٥,٣}$$

$$+ \frac{(٥ - ٤,٩)^{٢}}{٤,٩} + \frac{(٨ - ٣,٨)^{٢}}{٣,٨} = ٢١,٤٤$$

وحيث أن درجات الحرية = (د – ١) (ص – ١) = (٣ – ١) (٥ – ١)
= ٢ × ٤ = ٨

ومستوى الدلالة المطلوب هو ٠,٠٥ ، فإن قيمة مربع كاى النظرية من توزيعها فى الجدول هى ١٥,٥١ . وبما أن القيمة المحسوبة لمربع كاى (٢١,٤٤) أكبر من القيمة النظرية فإننا نرفض فرض العدم عند مستوى الدلالة ٠,٠٥ أى أننا نرفض القول بأنه لاتوجد اختلافات جوهرية بين درجات الانحدار للأنواع الثلاثة من التركيب الصخرى. أو بمعنى آخر فإننا نقبل الفرض البديل وهو أنه ليس من المحتمل أن يكون الاختلاف المشاهد بين درجات انحدار السفوح للأنواع الثلاثة من التركيب الصخرى فى المناطق الثلاث راجعاً للصدفة وحدها، بل هو إختلاف «حقيقى» وجوهرى يبرز أهميته تأثير اختلاف نوع التركيب الصخرى على درجات انحدار السفوح التى تكونت فوقه.

ثانياً: اختبار كولموجوروف – سميرنوف

Kolmogorov - Smirnov Test

(اختبار «د»، D - test)

يشبه هذا الاختبار إختبار مربع كاى فى أنه يستخدم لقياس مدى التطابق بين توزيعين أحدهما توزيع فعلى والآخر توزيع نظرى (احتمالى). ولكن يفضل اختبار كولموجوروف – سمير نوف على مربع كاى لأنه أسهل فى التطبيق، كما أن استخدامه لايتطلب شروطاً خاصاً مثلما يتطلب تطبيق اختبار مربع كاى. وفى كثير من الدراسات الجغرافية يتطلب التحليل الكمى التعرف أولاً على حقيقة ما إذا كانت بيانات عينة الدراسة تنتمى مثلاً لمجتمع تتمثل فيه خصائص نوع معين من التوزيعات الاحتمالية المعروفة مثل توزيع بواسون، أو التوزيع المعتدل (الطبيعى). فلو كان لدينا بيانات عن تكرار ظاهرة ما على فترات معلومة فإنه يمكن أن نضع فرضاً إحصائياً بعشوائية حدوث أو تكرار هذه الظاهرة فى تلك الفترة، ونقوم بحساب التوزيع النظرى المتوقع فيما لو كانت هذه البيانات تتوزع حسب توزيع بواسون أو تتوزعا توزيعاً معتدلاً. ثم يتم اختبار هل التوزيع الفعلى لبيانات هذه الظاهرة مطابق للتوزيع النظرى المتوقع بحساب قيمة «د» (أى الفرق بين احتمالات كل من التوزيعين الفعلى والمتوقع). ودرجات الحرية لقيمة «د» تساوى مجموع التكرارات الكلية لبيانات الظاهرة.

وسوف نوضح خطوات حساب قيمة «د» لاختبار كولموجوروف - سميرنوف لبيانات جدول التوزيع التكرارى لهبوب العواصف الرعدية فى كل سنة لفترة ١٠٠ سنة.

وبما أن هذه الظاهرة تعد من الظواهر النادرة الحدوث، فإن توزيعها التكرارى المشاهد يتطابق، بدرجة كبيرة، مع توزيع بواسون الاحتمالى (النظرى). وبحساب قيمة «د» يمكن معرفة هل الاختلاف بين التوزيع النظرى أو المتوقع (توزيع بواسون) والتوزيع الفعلى أو المشاهد (الحقيقى) يرجع إلى الصدفة أم هو اختلاف حقيقى. وخطوات حساب قيمة «د» موضحة بالجدول رقم (١٣ - ٥).

جدول رقم (١٠ - ٥)

التوزيع الفعلى (المشاهد) لهبوب العواصف الرعدية فى فترة ١٠٠ سنة والتوزيع النظرى (بواسون) لها

عدد العواصف فى كل سنة	عدد السنوات التى يتكرر فيها العواصف	الاحتمال		الاحتمال المتجمع		
		المشاهد	بواسون	المشاهد	بواسون	الفرق (د)
صفر	٣٠	٠,٢٠	٠,١٣	٠,٣٠	٠,١٣	٠,١٧
١	٢٥	٠,٢٥	٠,٢٦	٠,٥٥	٠,٣٩	٠,١٦
٢	١٥	٠,١٥	٠,٢٧	٠,٧٠	٠,٦٦	٠,٠٤
٣	١٢	٠,١٢	٠,١٩	٠,٨٢	٠,٨٥	٠,٠٣
٤	٥	٠,٠٥	٠,٠١٠	٠,٨٧	٠,٩٥	٠,٠٨
٥	٦	٠,٠٦	٠,٠٤	٠,٩٣	٠,٩٩	٠,٠٦
٦	٢	٠,٠٢	٠,٠١	٠,٩٥	١,٠٠	٠,٠٥
٧	١	٠,٠١	٠,٠	٠,٩٦	١,٠٠	٠,٠٤
٨	٠	٠,٠	٠,٠	٠,٩٦	١,٠٠	٠,٠٤
٩	٠	٠,٠	٠,٠	٠,٩٦	١,٠٠	٠,٠٤
١٠	٤	٠,٠٤	٠,٠	١,٠٠	١,٠٠	٠,٠٠
المجموع	١٠٠	١,٠٠	١,٠٠			

ويلاحظ من الجدول السابق أننا قمنا بحساب الاحتمال المشاهد للتوزيع وذلك بتحويل التكرارات المطلقة إلى احتمالات أو تكرارات نسبية عن طريق قسمة كل تكرار من التكرارات التى تحدث بها العواصف على المجموع الكلى لعدد العواصف. ولاختبار فرض عشوائية التوزيع، فإننا نقارن هذا التوزيع الاحتمالى المشاهد بالتوزيع النظرى أو توزيع بواسون الاحتمالى الذى يناسب مثل هذه الحالة من المقارنة والتى تفترض أن حدوث العواصف الرعدية يتكرر عشوائياً فى الفترة الزمنية (١٠٠ سنة) قيد الفحص. وتجرى المقارنة لمعرفة مدى تطابق أو اقتراب الاحتمالات المشاهدة من الاحتمالات النظرية باختبار جودة التوفيق Goodness of Fit بينهما. وفى هذا المجال يمكن استخدام أو تطبيق إختبار كولموجوروف – سميرنوف لأداء هذه المهمة.

وعند مقارنة توزيعين احتماليين بواسطة إختبار كولموجوروف – سميرنوف فإنه يجب تحويلهما إلى توزيعات احتمالية متجمعة Cumulative Propability distributions. ويكون فرض العدم لاختبار جودة التوفيق هو أن بيانات العينة التى ينتج عنها توزيعها احتمالياً فعلياً (مشاهداً) تكون لعينة مسحوبة من مجتمع يمتلك توزيعاً نظرياً خاصاً يشابه، فى هذه الحالة، توزيع بواسون الاحتمالى. وتحسب إحصائية الاختبار «د» على أساس أنها عبارة عن أقصى فرق (اختلاف) مطلق بين الاحتمال المتجمع لكل من التوزيعيين الفعلى (المشاهد) والنظرى. ومن الجدول السابق نجد أن أقصى فرق (أو قيمة إحصائية الاختبار «د») هو ٠,١٧.

وحيث أن قيمة «د» المحسوبة (٠,١٧) أكبر من القيمة النظرية (٠,١٤) من الجدول الخاص بها (أنظر ملحق الجداول الإحصائية فى نهاية الكتاب) بدرجات الحرية ١٠٠ (مجموع التكرارات) وفى مستوى دلالة ٠٥, فإننا نرفض فرض العدم عند مستوى الدلالة المعلوم. ويعنى هذا أن هناك إحتمال أقل ٠٥, (أو فى ٥ حالات فى كل ١٠٠ حالة) أن تمثل بيانات العينة توزيعاً له خصائص توزيع بواسون الاحتمالى. وتكون النتيجة النهائية بالنسبة للباحث الجغرافى هى أن التوزيع التكرارى الفعلى لحدوث العواصف الرعدية لايمكن أن يكون توزيعاً عشوائياً أو أن احتمال أن يكون كذلك ضعيف للغاية.

ثالثاً: اختبار مان – هويتنى

The Mann - Whitney Test

(اختبار «ى» U - test)

يعتبر إختبار مان – هويتنى من أبسط الأساليب الكمية غير البارامترية التى تبحث فى مقارنة مجموعتين من بيانات المعاينة لبيان ما إذا كانت هاتان المجموعتان مسحوبتين من مجتمعين مختلفين ولهما نفس المتوسط الحسابى أم مسحوبتين من مجتمع واحد، وهو بذلك يعد إختبار بديل لاختبار ستيودنت – ت. ولايشترط اختبار مان – هويتنى، مثله فى ذلك مثل بقية الأساليب اللابارامترية الأخرى، أن يكون توزيع البيانات لكل عينة توزيعاً متماثلاً (معتدلاً)، كما لايرتبط تطبيقية بنوع معين من أنواع البيانات ولكن يفضل استخدامه إذا كانت البيانات من نوع البيانات الترتيبية Ordinal الثنائية (المزدوجة) والمتساوية وغير المتساوية فى عدد مفرداتها.

ويستخدم إختبار مان – هويتنى (ى) لاختبار دلالة الفرق بين وسيطى عينتين أو اختبار فرض العدم القائل بأن العينتين مسحوبتان من مجتمع واحد وبالتالى يجب أن لايكون هناك اختلافاً جوهرياً أو حقيقياً بين بيانات العينتين، وأن أى اختلاف بينهما إنما يرجع الصدفة. وتحسب قيمة الاختبار (ى) من المعادلتين الآتيتين معا:

$$١- (ى) = ن_١ \times ن_٢ + \frac{ن_١ (ن_١ + ١)}{٢} - مجـ ت_١$$

$$٢- (ى) = ن_١ \times ن_٢ + \frac{ن_٢ (ن_٢ + ١)}{٢} - مجـ ت_٢$$

حيث $ن_١$ هى حجم العينة الأولى، $ن_٢$ هى حجم العينة الثانية، مجـ $ت_٢$ هى مجموع الرتب للعينة الأولى، مجـ $ت_٢$ هى مجموع الرتب للعينة الثانية.

وتوزيع مان – هويتنى غير منتظم ويقترب من التوزيع المعتدل كلما زادت أحجام العينات، لذلك فى الحالات التى يزداد فيها حجم العينات فإنه يمكن

استخدام توزيع (ز) – جداول المساحات تحت المنحنى المعتدل – للاستدلال على الفرض الموضوع للاختبار.

وسنوضح فى الأمثلة التالية خطوات حساب الفرق بين بيانات عينتين بواسطة اختبار مان–هويتنى (ى).

مثال (١):

لدراسة نسبة المحتوى المائى (نسبة الرطوبة) فى التربة أخذت عدة قياسات فى منطقتين زراعيتين فكانت نسبة الرطوبة فى التربة فى كل منها كما يلى:

المنطقة الأولى: ١٥,٢ ١٠,٧ ٨,٦ ٩,٣ ١٢,٤ ١١,١

المنطقة الثانية: ٨,٤ ٩,٣ ٨,٧ ١٠,٢ ١٠,٠

وفرض العدم فى هذه الحالة هو أن نسبة المحتوى المائى (نسبة الرطوبة) فى المنطقة الأولى أكبر من أو تساوى مثيلتها فى المنطقة الثانية بمستوى دلالة ٠,٠٥ ولاختبار الفرض السابق نحسب قيمة (ى) وذلك بافتراض أن صفة نسبة المحتوى المائى للتربة لاتتوزع توزيعاً معتدلاً، كما أن كل عينة مستقلة عن الأخرى وكلاً العينتين أخذت بطريقة عشوائية. وخطوات الحساب موضحة فى الجدول التالى:

جدول رقم (١٠-٦)

حساب قيمة (ى) من بيانات عينتين لنسبة المحتوى المائى فى التربة

العينة الأولى	الرتبــة	العينة الثانية	الرتبــة
١٥,٢	١٢	٨,٤	١
١٠,٧	٨	٩,٣	٤,٥
٨,٦	٢	٨,٧	٣
٩,٣	٤,٥	١٠,٢	٧
١٢,٤	١٠	١٠,٠	٦
١١,١	٩		
مجـ	٤٤,٥		٢١,٥
	$ن_1 = ٦$		$ن_2 = ٥$

من الجدول السابق يتضح أننا قمنا بترتيب جميع المفردات للعينتين فى ترتيب واحد، كما يلاحظ أننا أعطينا رتبة موحدة للمفردات المتساوية القيمة فمثلاً المفردة ٩,٣ تكررت مرتين فى البيانات، وبما أن موضعها يحتل الرتبة الرابعة والخامسة بين الترتيب الكلى، فإننا أعطينا لها الترتيب ٤,٥ (أى (٤ + ٥) ÷ ٢). وبالمثل إذا كان لدينا ثلاث مفردات متطابقة فى قيمها تحتل مواقعها الرتبة الخامسة والسادسة والسابعة فإننا نعطى لكل منها الرتبة ٦ (أى (٥ + ٦ + ٧) ÷ ٣)، وهكذا.
ولاختبار الفرق بين العينتين فى نسبة المحتوى المائى فى التربة نحسب قيمة (ى) باستخدام كل من المعادلتين السابقتين على النحو التالى:

$$(ى) = ٦ \times ٥ + \frac{٦\,(٦ + ١)}{٢} - ٤٤,٥ = ٣٠ + ٢١ - ٤٤,٥$$
$$= ٥١ - ٤٤,٥ = ٦,٥$$

$$(ى) = ٦ \times ٥ + \frac{٥\,(٥ + ١)}{٢} - ٢١,٥ = ٣٠ + ١٥ - ٢١,٥$$
$$= ٤٥ - ٢١,٥ = ٢٣,٥$$

ولاختبار معنوية قيمة (ى) نحتاج دائماً إلى أصغر قيمة من قيمتى (ى) المحسوبتين من المعادلتين السابقتين. ولتسهيل عملية الحساب فإنه يمكن الحصول على قيمة (ى) من أية معادلة من المعادلتين ومنها يمكن الحصول على القيمة الأخرى وذلك لأن مجموع القيمتين لابد وأن يساوى حاصل ضرب حجم العينة الأولى فى حجم العينة الثانية ($ن_١ \times ن_٢$). ففى المثال $ن_١ \times ن_٢ = ٣٠$، فإذا كانت قيمة (ى) الأولى تساوى ٦,٥ فإن قيمة (ى) الأخرى تساوى٢٣,٥ (أى ٣٠ − ٦,٥).

ونظراً لأن الفرض البديل فى هذه الحالة محدد الاتجاه فإننا نختار إختبار المعنوية من طرف واحد One - tailed test لقبول أو رفض فرض العدم بمستوى الدلالة المطلوب. وتكون النتيجة أنه تبعاً لأن قيمة (ى) الصغرى المحسوبة وهى ٦,٥ أكبر من قيمة (ى) النظرية التى تساوى ٥ عند مستوى الدلالة ٠٥, لاختبار الطرف الواحد وعندما تكون $ن_١ = ٦$، $ن_٢ = ٥$ (انظر ملحق الجداول الاحصائية)

فإننا نقبل فرض العدم، أى أن نسبة المحتوى المائى فى التربة (نسبة الرطوبة) فى المنطقة الأولى تساوى أو أكبر «جوهرياً» من مثيلتها فى المنطقة الثانية.

وتجدر الإشارة هنا إلى أن اختبار الدلالة (المعنوية) الخاص بإختبار مان-هويتنى (ى)، دون معظم اختبارات الدلالة الأخرى، يرفض فيه فرض العدم الموضوع للاختبار عند مستوى دلالة معين إذا كانت قيمة (ى) المحسوبة أقل من أو تساوى قيمة (ى) النظرية (القيمة الحرجة U-Critical Value).

مثال (٢):

فى دراسة لمعرفة اختلاف تأثير العوامل البحرية والنهرية على استدارة الرواسب الحصوية جمعت بيانات عن معامل الاستدارة[1] من عينتين عشوائيتين من الحصى: تمثل إحداهما مجتمع الرواسب على الأرصفة البحرية (الشواطئ المرتفعة) والأخرى تمثل مجتمع الرواسب الدلتاوية النهرية، فكانت معاملات استدارة الحصى فى كل عينة كما هو موضح بالجدول التالى؛ والمطلوب اختبار الفرض القائل بوجود إختلاف فى تأثير كل من العوامل البحرية والنهرية على استدارة الرواسب. الحصوية أو بعبارة أخرى هل يمكن الحكم على أن العينتين مسحوبتان من مجتمع واحد بمستوى دلالة ٠,٠٥.

(١) معامل الاستدارة هو مقياس كمى تقاس به استدارة الرواسب الحصوية، ويمكن حسابه من المعادلة:

$\frac{٢ \text{ نق}}{\text{ل}}$ × ١٠٠٠ حيث نق هى قيمة نصف قطر أصغر تحدب للحصوة، ل هى أكبر طول للحصوة

ويتراوح معامل الاستدارة للرواسب الحصوية ١ - ١٠٠٠ (الكرة المثالية)

جدول رقم (١٠ – ٧)

حساب قيمة (ى) من بيانات عينتين لاستدارة الرواسب الحصوية

معامل استدارة الحصى البحرى	الرتبـــة	معامل استدارة الحصى النهرى	الرتبـــة
٣٤٩	١٢	٣٢٥	٦
٢٥٨	١	٤٨٥	٢٣
٣٣٥	٧	٤٥٦	٢٢
٣١٧	٤	٤٣٨	٢٠
٣١٠	٣	٤٥٥	٢١
٣٠٤	٢	٣٨٨	٨
٣٧١	١٧	٣٦٩	١٦
٣٤٠	٥	٣٨٩	١٨
٣٤٤	١٠	٣٦٥	١٥
٣٤٥	١١	٣٥١	١٤
٤٠٥	١٩	٣٥٠	١٣
٣٠٤	٥		
$ن_١$ = ١٢	المجموع (مجـ $ت_١$) = ١٠٠	$ن_٢$ = ١١	المجموع (مجـ $ت_٢$) = ١٧٦

بعد إعطاء قيم معاملات الاستدارة فى العينتين الرتب الخاصة بها، وبعد جمع رتب كل عينة على حدة، نقوم بحساب قيمة (ى) على النحو التالى:

قيمة (ى) المحسوبة للعينة الأولى = $ن_١ \times ن_٢ + \frac{ن_١ (ن_١ + ١)}{٢} - مجـ\ ت_١$

$$= ١٢ \times ١١ + \frac{١٢ (١٢ + ١)}{٢} - ١٠٠$$

$$= ١٣٢ + ٧٨ - ١٠٠ = ١١٠$$

قيمة (ى) المحسوبة للعينة الثانية = $ن_١ \times ن_٢ + \frac{ن_٢ (ن_٢ + ١)}{٢} - مجـ\ ت_٢$

$$= ١٢ \times ١١ + \frac{١١ (١١ + ١)}{٢} - ١٠٠$$

$$= ١٣٢ + ٧٨ - ١٠٠ = ١١٠$$

وبطبيعة الحال فإننا نحتاج إلى القيمة الأصغر من قيمتى (ى) المحسوبتين وهى (٢٢) للاستدلال عن معنوية الفرض الموضوع للاختبار. وبمقارنة قيمة (ى) المحسوبة (٢٢) بقيمة (ى) النظرية فى جدول الاختبار من طرف واحد فى المستوى دلالة ٠٥, ولحجم العينة ١٢ ، ١١ وهى (٣٨) فإنه يمكن فى هذه الحالة رفض فرض العدم القائل أن العينتين مسحوبتان من مجتمع واحد، أى أن هناك تأثير للعوامل البحرية والنهرية، كل على حدة، على استدارة الرواسب الحصوية بنوعيها - يظهر فى وجود فروق جوهرية بين معامل استدارة الرواسب الحصوية فى العينتين.

مثال (٣):

إذا كانت لدنيا عينتين عشوائيتين تمثل كل منهما إنتاج فول الصويا (طن / فدان) لعدد من الأحواض الزراعية فى منطقتين وذلك بغرض إختبار أن كل من متوسطى العينتين متساويين أو بمعنى آخر هل العينتان مأخوذتان من مجتمعين مختلفين أم من مجتمع واحد؟ وفى حالة عدم تساوى المتوسطين فإننا نتوقع أن متوسط العينة الأولى يزيد عن متوسط العينة الثانية، أى أن الاختبار المناسب هو إختبار (ى) فى مجموعتين من طرف واحد بمستوى دلالة ٠,٠٥ وكانت بيانات العينتين موضحة فى الجدول التالى:

جدول رقم (١٠ - ٨)

إنتاج فول الصويا (طن/ فدان) في منطقتين مرتبة حسب مقدار الإنتاج

الإنتاج فى المنطقة الأولى	الرتبـة (ت)	الإنتاج فى المنطقة الثانية	الرتبـة (ت)
٣,٢٨	١٣	٣,٠٥	٦
٣,١٩	١١	٢,٨٦	٢
٣,٦٣	٢٣	٣,١٦	٩
٣,٥٦	٢٠	٣,١٧	١٠
٣,٦١	٢١	٣,٢٠	١٢
٢,٩٨	٤	٣,٣٣	١٥
٣,٤٥	١٨	٢,٩٦	٣
٣,٤٠	١٧	٣,٢٩	١٤
٣,٤٩	١٩	٣,١٥	٨
٣,٣٤	١٦	٣,٠١	٥
		٣,٦٤	٢٢
		٣,١٤	٧
		٢,٨٢	١
المجموع = ن_1 = ١٠	١٦٢	ن_2 = ١٣	١١٤

لاختبار الفرض السابق نحسب قيمة (ى) بالمعادلتين السابقتين كما يلى:

$$(\text{ى}) = \text{ن}_1 \times \text{ن}_2 + \frac{\text{ن}_1 (\text{ن}_1 + 1)}{2} - \text{مجـ ت}_1$$

$$= 12 \times 11 + \frac{10 (10 + 1)}{2} - 162$$

$$= 130 + 55 - 162 = 23$$

$$(\text{ى}) = \text{ن}_1 \times \text{ن}_2 + \frac{\text{ن}_1 (\text{ن}_2 + 1)}{2} - \text{مجـ ت}_1$$

$$= ١٢ \times ١١ + \frac{١٣ (١٣ + ١)}{٢} - ١١٤$$

$$= ١٣٠ + ٩١ - ١١٤ = ١٠٧$$

وحيث أن قيمة (ى) الصغرى المحسوبة وهى (٢٣) أقل من قيمة (ى) النظرية لحجم العينة ١٠، ١٣ بمستوى دلالة ٠٥, وفى اختبار الطرف الواحد وهى (٣٧) فإننا نرفض فرض العدم ونقبل الفرض البديل وهو أن متوسط إنتاج فول الصويا فى الأحواض الزراعية فى المنطقة الأولى (العينة الأولى) يزيد عن متوسط الإنتاج للأحواض الزراعية فى المنطقة الثانية (العينة الثانية) بمستوى دلالة ٠,٠٥ والنتيجة التى يمكن أن نستخلصها، اعتماداً على البيانات السابقة تؤكد تأثير العوامل الجغرافية فى الإنتاج الزراعى بعامة وإنتاج فول الصويا بخاصة فى المنطقتين.

رابعاً: اختبار ويلكوكسون

Wilcoxon Test

(اختبار «ق»)

يعد إختبار ويلكوكسون من أبسط أساليب المقارنة الاحصائية اللاباراميترية التى يمكن استخدامها لاختبار دلالة الفروق (الاختلافات) بين رتب أزواج القيم (المفردات) المتماثلة للمتغير موضع الاختبار. ويتطلب الاختبار إذن أن تكون البيانات الأصلية من نوع بيانات الفترة ولكن جدولتها تكون على شكل رتب، كما يعتمد مستوى الدلالة (المعنوية) المستخدم لرفض فرض العدم لهذا الاختبار على اتجاه ومقدار الاختلاف بين الرتب لكل زوج من أزواج القيم المتماثلة (أى عندما تكون إحدى القيم (أ) أكبر من القيمة الأخرى (ب) أو العكس). كما يشترط عند استخدام إختبار ويلكوكسون لأزواج المفردات المتماثلة فى العدد أن

تكون مفردات كل زوج متجانسة وتخصص إحدى المفردات من كل زوج للعينة الأولى بينما تخصص المفردة الأخرى للعينة الثانية.

ولايقتصر تطبيق إختبار ويلكوكسون على الحالات التى يكون فيها حجم العينة صغيراً (ن أقل من ٣٠) فحسب، بل أنه يستخدم كذلك فى الحالات التى يكون فيها حجم العينة كبيراً (ن أكبر من ٣٠)، ولاتختلف عملية الاختبار فى الحالتين إلا فى الخطوات الحسابية الأخيرة. ويتلخص الاختبار فى كل حالة فى اعتبار الفروق بين رتبتى كل زوج من القيم عينة عشوائية من مجتمع تمثل جميع الفروق الممكنة، وإختبار هل العينتان مأخوذتان من مجتمع واحد أم من مجتمعين لهما نفس المتوسط الحسابى أو مسحوبتان من مجتمعين مختلفين.

ولتوضيح طريقة مقارنة بيانات **العينات الصغيرة الحجم** (ن أقل من ٣٠) نذكر المثال التالى:

مثال:

لمقارنة درجات الانحدار على جانبى أحد الأودية النهرية (باتجاه شمالى غربى «ش. غ» لأحد الجوانب واتجاه جنوبى شرقى «جـ. ق» للجانب الآخر) أخذ القياسات الموضحة فى الجدول التالى (جدول رقم ١٠ - ٩). وتمثل القياسات المزدوجة والمتماثلة فى العدد بالجدول المسافة الكنتورية (بالمليمتر) المأخوذة من قطاعات عرضية ذات أبعاد متساوية فيما بينها على خريطة كنتورية بمقياس كبير. وكما نعلم أن هناك علاقة وطيدة بين المسافة الكنتورية ودرجة الانحدار، فكلما اتسعت المسافة بين خطوط الكنتور كلما دل ذلك على ضعف انحدار سطح الأرض والعكس صحيح. والمطلوب هو إختبار ما إذا كان الفرق بين متوسطى درجات الإنحدار لجانبى الوادى يمثل فرقاً جوهرياً (حقيقيا) أو أنه فرق يرجع إلى الصدفة المطلقة نتيجة خطأ المعاينة. وبعبارة أخرى فإن فرض العدم المطلوب اختباره هو أن مقدار الانحدار لايختلف على منحدرات جانبى الوادى فى مقابل الفرض البديل أن مقدار الانحدار يختلف ويتنوع على جانبى الوادى من مكان لآخر حسب درجة التوجيه Aspect وذلك بمستوى دلالة ٠,٠٥.

جدول رقم (١٠ - ٩)

طريقة حساب إختبار ويلكوكسون من بيانات الفترة الكنتورية (بالملليمتر) على جانبى أحد الأودية النهرية

الجانب المواجه للشمال الغربى (أ)	الجانب المواجه للجنوب الشرقى (ب)	الفرق |أ - ب|	رتب الفرق	الرتب الخاصة	
				أ > ب	أ < ب
١٠	١١	١	١		١
٧	٧	صفر	تحذف من الترتيب لان أ = ب		
٩	١٣	٤	٦		٦
١٠	٨	٢	٢,٥	٢,٥	
١٥	١٢	٣	٤,٥	٤,٥	
١٢	١٥	٣	٤,٥		٤,٥
٨	١٠	٢	٢,٥		٢,٥
٩	١٥	٦	٧		٧
المجموع				$ر_1$ = ٧	$ر_2$ = ٢١

١- الفروض: H_0: متوسط انحدار السفوح الشمالية الغربية = متوسط انحدار السفوح الجنوبية الشرقية.

H_1: متوسط انحدار السفوح الشمالية الغربية ≠ متوسط انحدار السفوح الجنوبية الشرقية.

٢- مستوى المعنوية (α) = ٠,٠٥ والاختبار فى هذه الحالة من طرفين تبعاً لافتراض تساوى المتوسطين، ولعدم تحديد أى من الجانبين أشد انحداراً من الآخر. وبذلك إذا كان احتمال حدوث هذه البيانات بالصدفة تحت شروط الاختبار أقل من ٢٠٠١ فإن نرفض فرض العدم، أى أن الفرضية تكون فى هذه الحالة غير صحيحة.

٣- بعد حساب الفرق بين أزواج القيم (مع إهمال الإشارة) تعطى الفروق رتباً تختلف حسب مقدار الفرق. ويعطى للفروق المتساوية متوسط الرتب الطبيعية لهذه الفروق. فمثلاً الفروق ٢، ٢ رتبتهما على الترتيب هما ٢ ، ٣ ومتوسطهما هو ٢,٥، وبالمثل الفروق ٣، ٣ رتبتهما على الترتيب هما ٤، ٥ وتكون الرتبة المتوسطة لهما هى ٤,٥. ثم نفصل بين ترتيب (رتب) الفروق المحسوبة بين العينتين الأولى (أ) و (ب)، أى عندما تكون قيم العينة الأولى أكبر من القيم فى العينة الثانية (ب)، والعكس عندما تكون قيم العينة الثانية (ب) أكبر من العينة الأولى (أ). ويكون أقل مجموع لهذه الرتب هو القيمة المختبرة، والتى يرمز لها بالرمز (ق)، وذلك على أساس أنه إذا كان مجموع الرتب لكل من $ر_١$، $ر_٢$ متساوياً (أى أن $ر_١ - ر_٢ =$ صفر) فإن فرض العدم يكون صحيحاً طالما أن الفرق بين أزواج القيم يتوزع عشوائياً بين القطاعات التى تكون فيها مفردات العينة (أ) أكبر من مفردات العينة (ب) وبالمثل عندما تكون (ب) أكبر من (أ). أما إذا كان الفرق بين كل من $ر_١$، $ر_٢$ أكبر من الصفر فإن ذلك يتخذ دليلاً كلما دل ذلك على عدم صحة فرض العدم. وحيث أن مجموع الرتب الخاصة ($ر_١ + ر_٢$) يكون ثابتاً لعينة حجمها (ن) فإن هذا المجموع يمكن وضعه فى الصيغة التالية:

$$\frac{١}{٢} ن (ن + ١)$$

وبذلك فإنه كلما زاد الفرق بين $ر_١$، $ر_٢$ كلما قلت القيمة الصغرى لهما وكلما دل ذلك على عدم صحة فرض العدم وبالتالى يمكن رفضه.

٤- منطقة الرفض: بالرجوع إلى جدول اختبار ويلكوكسون (راجع ملحق الجداول الاحصائية) والتى تتضمن القيم الحرجة (ق) للعينات المختلفة الحجم: من ن = ٦ حتى ن = ٢٥ (إذ لايمكن تطبيق اختبار ويلكوكسون على العينات ذات الحجم أقل من ٦ أزواج من المفردات يمكن استخراج القيمة الحرجة التى نقارن بها القيمة المحسوبة (ق). وحيث أن مستوى الدلالة المطلوب هو ٠,٠٥ وأن الفرض البديل غير محدد الاتجاه فإن الاختبار للمثال بين أيدينا من نوع الاختبار من الطرفين. وطبقاً للأسس الموضوعة سابقاً فإذا كانت قيمة (ق) المحسوبة أقل من

القيمة الحرجة المناظرة لها بالنسبة لحجم العينة قيد الاختبار فإننا نرفض فرض العدم ونستنتج أن الفرق بين العينتين هو فرق جوهرى أو حقيقى عند مستوى الدلالة ٠,٠٥ .

٥- الاستنتاج: وبما أن القيمة الحرجة لهذا المثال هى ٢ عند ن = ٧ (بعد أن أسقطنا من حسابنا أحد أزواج القيم من الحجم الأصلى للعينة لعدم وجود فرق بين قيمتيه) والقيمة المحسوبة (ق) هى ٧، فإن بيانات العينة لاتقدم دليلاً كافياً لرفض فرض العدم. أو بعبارة أخرى أنه تحت شروط الاختبار نجد أن قيمة (ق) المحسوبة أكبر من القيمة الحرجة بمستوى معنوية ٠٥ ,٠، وبالتالى نقبل فرض العدم ونستنتج أن هذا الفرق يرجع إلى الصدفة المطلقة. ويعنى قبولنا لفرض العدم فى هذه الحالة أنه لاتوجد اختلافات جوهرية بين درجات الانحدار على جانبى هذا الوادى النهرى تبعاً لاختلاف درجة التوجيه Aspect، مما يقلل من دور هذا العامل البيئى فى التأثير على درجات الانحدار للمنحدرات فى حوض هذا النهر.

وفى حالة العينات الكبيرة (ن أكبر من ٣٠) يصمم اختبار ويلكوكسون بنفس الطريقة التى اتبعناها سابقاً لاختبار الفروض التى تتعلق بالمقارنة بين المتوسطات لعينات صغيرة أقل من ٣٠ من أزواج القيم المتماثلة فى العدد، فيما عدا أن الاستنتاج (أى قرار قبول أو رفض فرض العدم) فى هذه الحالة يختلف قليلاً عن مثيله للعينات الصغيرة على أساس أن الفرق بين الرتب يدخل فى الاعتبار عند تقرير ما إذا كان صغر قيمة (ق) كافياً لتأكيد رفض فرض العدم. ونظراً لأن جدول القيم الحرجة لاختبار ويلكوكسون لايشتمل على قيم حرجة لمجموع الفرق بين أكثر من ٢٥ زوجاً من قيم العينات، فإنه فى حالة العينات الكبيرة يؤول توزيع القيمة (ق) تحت شروط الاختبار إلى الطرف الأيسر لمنحنى التوزيع المعتدل الذى متوسطه الحسابى يساوى نصف المجموع الكلى للرتب [أى يساوى $\frac{1}{4}$ن(ن+١)]،

وانحرافه المعيارى يساوى $\sqrt{\frac{[ن (ن + ١) (٢ ن + ١)]}{٢٤}}$. وبالتالى فإن احتمال

(ح) بأن تكون القيمة (ق) - تحت شروط الاختبار وفرض العدم - أقل من أو تساوى القيمة المحسوبة يمكن الحصول عليه بواسطة:

أ- تحويل الإحصائية (ق) إلى قيمة معيارية من قيم الإحصائية (ز) Z - Seore: عن طريقة ايجاد الفرق بينهما وبين متوسط توزيع المعاينة وقسمة هذا الفرق على الانحراف المعيارى لتوزيع المعاينة، أى أن:

$$\text{ز} = \frac{\text{ق} - \frac{1}{4}\,\text{ن}\,(\text{ن} + 1)}{\sqrt{\frac{[\text{ن}\,(\text{ن} + 1)\,(2\,\text{ن} + 1)]}{24}}}$$

ب- ايجاد قيمة الاحتمال (ح) المقابلة للقيمة المحسوبة من جداول التوزيع المعتدل المعيارى حسب نوع الاختبار (من طرف واحد أو من طرفين)، فإذا كانت قيمة الاحتمال المحسوبة أقل من قيمة احتمال مستوى الدلالة (α) فإننا نرفض فرض العدم ونقبل الفرض البديل، والعكس إذا كان احتمال القيمة المعيارية أكبر من احتمال مستوى الدلالة (α) فإن فرض العدم لايمكن رفضه.

ولتوضيح استخدام إختبار ويلكوكسون فى حالة العينات الكبيرة نذكر المثال التالى:

مثال:

الجدول رقم (١٠ - ١٠) يشتمل على بيانات لأعداد القوى العاملة (بالمئات) بمجموعة من القرى فى عام ١٩٧٠، ١٩٧٦. والمطلوب عن طريق استخدام اختبار ويلكوكسون تحديد احتمال أن مثل هذا الفرق الكبير بين مفردات العينتين، كما يبدو فى الجدول، يمكن حدوثه حتى إذا كانت مجتمعاتهما النظرية متشابهة ومتساوية فى معالمها.

جدول رقم (١٠ – ١٠): أعداد القوى العاملة (بالمئات) في مجموعة من القرى لعامي ١٩٧٠، ١٩٧٦

أ	ب	|أ – ب|	رتب الفرق	الرتب الخاصة		أ	ب	|أ – ب|	رتب الفرق	الرتب الخاصة	
				أ > ب	أ < ب					أ > ب	أ < ب
٨٦	١٠٢	١٦	١٣		١٣	٩٧	١٢٨	٣١	٢٢,٥		٢٢,٥
٦٤	٩٢	٢٨	٢٠,٥		٢٠,٥	٣٠	٤٠	١٠	٧		٧
٥٨	٥٩	١٢٥	١		١	٩٢	١٢٣	٣١	٢٢,٥		٢٢,٥
٧٣	٩٨	١	١٨		١٨	٩٨	١٥٥	٥٧	٢٩		٢٩
٢٥	٣٠	٥	٣,٥		٣,٥	٣٤	٢٦	٨	٦	٦	
٦٨	٨٨	٢٠	١٦		١٦	٢٥	٨٧	٦٢	٣٠		٣٠
٨٥	١٢٣	٣٨	٢٦		٢٦	٦٧	٥٠	١٧	١٤	١٤	
٥٩	٨٠	٢١	١٧		١٧	٧٠	٥٢	١٨	١٥	١٥	
١١٥	١١٠	٥	٣,٥	٣,٥		١١٣	٨٦	٢٧	١٩	١٩	
٣٦	٨٤	٤٨	٢٨		٢٨	٦٠	٧٤	١٤	١٠,٥		١٠,٥
٥٥	٤٠	١٥	١٢	١٢		٦٦	١٠١	٣٥	٢٥		٢٥
٦٩	٦٥	٤	٢	٢		٨٩	١٣٢	٤٣	٢٧		٢٧
٣٥	٤٦	١١	٨		٨	٥٧	٤٥	١٢	٩	٩	
٧٨	١١٢	٣٤	٢٤		٢٤	٤٣	٧٢	٢٨	٢٠,٥		٢٠,٥
٧٦	٦٩	٧	٥	٥		٨	٢٢	١٤	١٠,٥		١٠,٥
المجموع										ر١ = ٨٥,٥	

(أ = العمالة فى ١٩٧٠، ب = العمالة فى ١٩٧٦)

وتكون خطوات الاختبار كما يلى:

١- H_0: لايوجد اختلاف بين أعداد القوى العاملة لعامى ١٩٧٠ و ١٩٧٦.

H_1: يوجد اختلاف بين أعداد القوى العاملة لعامى ١٩٧٠ و ١٩٧٦

٢- مستوى الدلالة (المعنوية α = ٠٥,٠

٣- باتباع نفس الإجراءات السابقة لاختبار العينات الصغيرة نحسب قيمة (ق) من واقع البيانات المشاهدة، وهى لهذا المثال، تساوى ٨٥,٥.

٤- منطقة الرفض: تحول قيمة (ق) المحسوبة إلى قيمة معيارية من قيم (ز) المعيارية كما يلى:

(أ) $$ز = \frac{ق - \frac{1}{4} ن (ن + 1)}{\sqrt{\frac{[ن (ن + 1) (2ن + 1)]}{24}}}$$

$$= \frac{85.5 - \frac{1}{4} \times 30 \times 31}{\sqrt{\frac{[30 \times 31 \times 61]}{24}}} = \frac{147}{49} = 3$$

(ب) نرجع إلى جدول التوزيع المعتدل المعيارى للحصول على قيمة (ز) النظرية. ونظراً لأن الفرض البديل غير محدد الاتجاه فإن الاختبار يكون من طرفين، وبالتالى فإن الاحتمال المقابل للقيمة المحسوبة (ز = ٣) هو ٠٠٣,٠.

٥- الاستنتاج: تحت شروط الاختبار نجد أن احتمال قيمة (ز) المحسوبة من القيمة (ق) وهو ٠٠٣,٠ وأقل بكثير من مستوى الرفض ٠٥,٠ لذا لايمكن أن يرجع هذا الفرق للصدف المطلقة، ونرفض فرض العدم القائل بأنه لايوجد فرق معنوى بين أعداد القوى العاملة فى عام ١٩٧٠ و ١٩٧٦، أى أن البيانات تمثل إختلافاً حقيقياً فى عدد القوى العاملة بين السنتين. وأكثر ما يمكن تفسيره جغرافياً من

هذا المثال هو أنه قد يوضح أن الاختلاف فى أعداد القوى العاملة فى العينتين (١٩٧٠ ، ١٩٧٦) قد يرجع إلى عامل زيادة الحركة فى مجتمع القوى العاملة أو عامل انتشار الصناعة في المناطق الريفية أو إلى العاملين معاً. ولكن لايمكن تأكيد هذه العوامل إلا عن طريق البحث المستفيض والدراسة التفصيلية وذلك لأن الأساليب الكمية كما نعلم، تثير كثيراً من التساؤلات أكثر من الإجابة عليها.

خامساً: اختبار كروسكال – واليس

The Kruskal - Wallis Test

(اختبار «هـ» H - test)

يستخدم إختبار كروسكال – واليس فى حالة المقارنة وبيان مدى التجانس أو الاختلاف بين ثلاث عينات أو أكثر ومدى صلتها بالمجتمع الأصلى الذى تمثله. ولذا فإنه يعتبر اختباراً يصلح كبديل لتحليل التباين، طالما أنه لايشترط، مثل غيره من أساليب المقارنة اللابارامترية، أن يكون توزيع مفردات أو قيم العينات متصفاً بصفة الاعتدالية، بل يتطلب فقط أن تكون البيانات من نوع البيانات الترتيبية (الرتب). وعلى الرغم من بساطة الاختبار وسهولة حسابه إلا أنه كأداة للتحليل لم يحظ بالاهتمام والاستخدام، حتى وقت قريب، فى مجال البحث الجغرافى.

ويعتمد تحليل الاختلافات بين العينات بإختبار «هـ» على فرض العدم القائل بأن العينات قد أخذت من مجتمعات لها توزيعات متطابقة، وأن أى اختلاف فيما بينها إنما يرجع إلى الصدفة المطلقة التى تتضمنها عملية المعاينة العشوائية. ويكون الفرض البديل المقابل لفرض العدم فى هذه الحالة هو أن الاختلاف بين العينات يعكس الاختلافات بين توزيع المجتمعات التى سحبت منها. ويمكن إختبار الاحصائية «هـ» بحسابها من المعادلة الآتية:

$$\text{هـ} = \frac{12}{\text{ن}(\text{ن}+1)} \times \text{مجـ} \frac{\text{ر}^2}{\text{ن}} - 3(\text{ن}+1)$$

حيث ن هى المجموع الكلى للمفردات فى كل العينات، ر هى مجموع الرتب لكل عينة، ن هى عدد المفردات فى كل عينة على حدة، مجـ $\frac{ر^{٢}}{ن}$ هو المجموع الكلى لمربعات مجموع الرتب المقسومة على عدد مفردات كل عينة على حدة. وهناك جداول محسوبة (راجع ملحق الجداول الإحصائية) لقيم «هـ» باحتمالات مختلفة لعينات ثلاث فقط بكل منها عدد من المفردات يتراوح بين مفردة واحدة وخمس مفردات. أما إذا زاد عدد المفردات عن خمس مفردات فإن قيمة «هـ» يكون لها توزيع احتمالى يتطابق مع التوزيع الاحتمالى لمربع كاى، بدرجات الحرية (ن – ١) حيث ن هى عدد العينات. ولذا لايقتصر تطبيق اختبار «هـ» على العينات الصغيرة فقط ولكن يمكن أيضاً استخدامه لتحليل بيانات العينات الكبيرة. ونقصد بالعينات الصغيرة فى هذا المقام بأنها العينات التى لايزيد عددها عن ثلاث عينات بكل منها عدد من المفردات (أو القياسات «ن») لايزيد عن ٥ مفردات.

والمثال التالى يوضح خطوات حساب الاختبار للعينات الصغيرة (ق =٣، ن <٥).

مثال (٢):

الجدول التالى (جدول رقم: ١٠– ١١) يوضح درجات الانحدار للسفوح الناجمة عن فعل الحركة البطيئة للمواد التى تتألف رواسبها من المفتتات الصخرية غير المغطاة بغطاءات نباتية Unvegetated Scree slopes التى تفككت من ثلاثة أنواع مختلفة من التكوينات الصخرية.

جدول رقم (١٠ - ١١)

درجات الانحدار لسفوح المفتتات الصخرية حسب نوع الصخور التى تفككت منها

حجر جيرى		شســـت		حجر جيرى	
درجة الانحدار	الرتبــة	درجة الانحدار	الرتبــة	درجة الانحدار	الرتبــة
١٧	١	٢١	٥	٣٠	١٢
٢٠	٤	١٨	٢	٢٦	١٠
١٩	٣	٢٥	٩	٢٢	٦
٢٤	٨	٢٨	١١		
٢٣	٧	..	..		
المجموع $ن_١$ = ٥	$ر_١$ = ٢٣	$ن_٢$ = ٤	$ر_٢$ = ٢٧	$ن_٣$ = ٣	$ر_٣$ = ٢٨

١- الفروض: H_0: لايوجد اختلاف بين درجات الانحدار لسفوح المفتتات التى تكونت من أنواع مختلفة من التكوينات الصخرية، أو بعبارة أخرى أن هناك احتمالاً كبيراً أن ترجع الاختلافات المشاهدة فى درجات الانحدار بين العينات الثلاث إلى الصدفة.

H_1 : أن الاختلافات فى درجة الانحدار بين العينات تعكس الاختلانات بين المجتمعات التى تمثلها هذه العينات، أى أنها إختلافات جوهرية حقيقية لاترجع إلى الصدفة.

٢- مستوى الدلالة (α) = ٠٥,٠.

٣- يمكن وضع الاختبار على الصورة الآتية:

$$هـ = \frac{١٢}{ن(ن+١)} \times مجـ \frac{ر^٢}{ن} - ٣(ن+١)$$

٤- هذه القيمة لها توزيع احتمالى هـ حسب عدد المفردات فى كل عينة، أى:

$ن_{١}$ = ٥ ، $ن_{٢}$ = ٤ ، $ن_{٣}$ = ٣

٥- منطقة الرفض: على أساس الشروط السابقة من (١) إلى (٢) نجد أن القيمة الحرجة لاختبار «هـ» التى تحدد الرفض هى ٥,٦٣١ ونقبل فرض العدم إذا كانت قيمة «هـ» المحسوبة أقل من ٥,٦٣١ أو لتكون قيمة هـ المحسوبة لها دلالة إحصائية لابد أن تساوى أو تزيد عن القيمة النظرية المقابلة لها.

٦- حساب قيمة «هـ» من واقع البيانات المشاهد، حيث:

$ر_{١}$ = ٢٣ ، $ر_{٢}$ = ١٧ ، $ر_{٣}$ = ٢٨ ، ن = ١٢

$ن_{١}$ = ٥ ، $ن_{٢}$ = ٤ ، $ن_{٣}$ = ٣ ، ق = ٣

$$\text{«هـ»} = \frac{١٢}{١٢\,(١٢ + ١)} \times \left(\frac{(٢٣)^{٢}}{٥} + \frac{(٢٧)^{٢}}{٤} + \frac{(٢٨)^{٢}}{٤}\right) - ٣\,(١٢ + ١)$$

$$= ,٠٧٦٩ \times (١٠٥,٨٠ + ١٨٢,٢٥ + ٢٦١,٣٣) - ٢٩$$

$$= (,٠٧٦٩ \times ٥٤٩,٣٨٣٣) - ٣٩$$

$$= ٣,٢٤٧$$

٧- الاستنتاج قيمة «هـ» المحسوبة من البيانات المشاهدة تقل عن القيمة الحرجة (٥,٦٣١)، ولذلك لايمكن رفض فرض العدم القائل بأنه لايوجد فرق جوهرى بين درجات الانحدار لسفوح المفتتات الصخرية التى تفككت من أنواع مختلفة من التكوينات الصخرية وذلك بمستوى دلالة ٠,٠٥ .

ويتبع نفس الأسلوب السابق عند تحليل بيانات العينات الكبيرة (ن أقل من أو تساوى ٣، ن أكبر من ٥) كما يتضح من المثالين التاليين.

مثال (٢):

الجدول التالى (جدول رقم ١٠ - ١٢) يشتمل على بيانات لعينة من الزوار لثلاث من الحدائق العامة فى أحد الأيام ممثلة فى المسافات (بالكيلو مترات) التى

قطعها زوار كل حديقة. والمطلوب اختبار فرض العدم القائل بأن العينات الثلاث للزوار تمثل مجتمعاً واحد للزوار أو مجتمعات متطابقة فى مقابل الفرض البديل القائل بأن هذه العينات تمثل مجتمعات مختلفة من الزوار وذلك بمستوى دلالة ٠٥,٠.

جدول رقم (١٠ - ١٢)

طريقة حساب إختبار «هـ» للمسافات (بالكيلو متر) التى قطعتها ثلاث عينات من الزوار لثلاث من الحدائق العامة

الحديقة (أ)		الحديقة (ب)		الحديقة (حـ)	
المسافة	الرتبة	المسافة	الرتبة	المسافة	الرتبة
٢٣	٧	١١	٣	١١٠	١٨
١٥	٤	١٧	٥	٨٤	١٦
٤٢	١٢	٣١	١١	٨٥	١٧
٨	١	٢٧	٩,٥	٤٥	١٣
١٠	٢	٦٣	١٤	٦٤	١٥
١٨	٦	٢٧	٩,٥	٢٥	٨
ن١ = ٦	ر١ = ٣٢	نب = ٦	ر٢ = ٥٢	نحـ = ٦	ر٣ = ٨٧

(١) الفروض H_0: لايوجد اختلاف بين المسافات للعينات الثلاث، أو بعبارة أخرى أن هذه العينات تمثل نفس المجتمع أو أنها تمثل ثلاث مجتمعات لها نفس المعالم.

H_1 : أن الفرق بين المسافات فى العينات الثلاث هو فرق جوهرى، أى أنها تمثل مجتمعات مختلفة.

(٢) مستوى الدلالة α = ٠٥,٠

(٣) يكون وضع الاختبار فى الصورة الاحصائية

$$هـ = \frac{١٢}{ن (ن + ١)} \times مجـ_ق \frac{ر^{٢}}{ن} - ٣ (ن + ١)$$

(٤) حيث أن: عدد العينات = ٣، وعدد المفردات (القياسات) فى كل عينة أكثر من ٥ مفردات فإن قيمة «هـ» لها توزيع احتمالى يتطابق مع توزيع مربع كاى بدرجات حرية ن - ١ حيث ن هي عدد العينات. ودرجات الحرية لهذا المثال = ٣ - ١ = ٢.

(٥) منطقة الرفض: على أساس ما سبق نجد أن قيمة «هـ» الحرجة التى تحدد منطقة الرفض من جدول توزيع كاى هى ٥,٩٩ بدرجات الحرية ٢ ونرفض H_0 فقط عندما تكون قيمة «هـ» المحسوبة من البيانات المشاهدة أكبر من أو تساوى القيمة ٥,٩٩.

(٦) حساب قيمة «هـ» من واقع البيانات المشاهدة: ويتم ذلك عن طريق ترتيب البيانات للعينات الثلاث مجتمعة ترتيباً تصاعدياً أو تنازلياً (من الأقل إلى الأعلى أو من الأعلى إلى الأقل) أى إعطاء كل مفردة من المفردات رتبة خاصة بها، وفى حالة تطابق بعض المفردات يعطى لكل منها رتبة تساوى متوسط رتب هذه المفردات. فمثلاً المفردة ٢٧ تكررت مرتين وأعطيت لها الرتبة ٩,٥ بدلاً من الرتبتين ٩ ، ١٠ (أى ٩ + ١٠ ÷ ٢ = ٩,٥). وبعد إتمام عملية ترتيب المفردات تجمع رتب مفردات كل عينة على حدة، ثم يربع مجموع رتب كل عينة ويقسم على عدد مفرداتها وأخيراً يجمع خارج عمليات القسمة للعينات فتحصل على مجموع متوسطات مربعات الرتب للعينات.

$$«هـ» = \frac{١٢}{١٨ \times ١٩} \times \left(\frac{(٣٢)^{٢}}{٦} + \frac{(٥٢)^{٢}}{٦} + \frac{(٨٧)^{٢}}{٦} \right) - ٣ \times ١٩$$

$$= ٠,٠٣٥ \times (١٧٠,٦٦ + ٤٥٠,٦٦ + ١٢٦١,٥) - ٥٧$$

= (٠,٠٣٥ × ١٨٨٢,٨٢) − ٥٧

= ٨,٩٠

(٧) الاستنتاج: قيمة «هـ» المحسوبة من البيانات المشاهدة أكبر من مثيلتها النظرية (٥,٩٩) بدرجات الحرية ٢ ولذلك نرفض فرض العدم القائل بأنه لاتوجد فروق جوهرية بين المسافات (بالكيلو متر) التى قطعها زوار الحدائق الثلاث، ونستنتج أن هناك فروقاً جوهرية بين العينات الثلاث من المسافات بما يدل على أن هذه العينات تمثل إختلافات حقيقة بين المجتمعات التى أتى منها الزوار، وذلك بمستوى دلالة ٠,٠٥ .

مثال (٣):

لدراسة خصائص حجم الرواسب الشاطئية وبيان ما إذا كانت طبيعة هذه الرواسب تدل على أنها من أصل مشترك، أو تؤدى إلى التعرف على العامل المرسب وعلى الظروف الجغرافية التى كانت سائدة أثناء الإرساب. وقد اختيرت عشوائياً لذلك ستة من الشواطئ وأخذ من كل منها عينة من الرواسب (فى منتصف المسافة تقريباً من المنطقة التى تتعرض لمتوسط المد والجزر) . وبعد تجفيف العينات الست وتعريضها لعملية النخل الجاف الميكانيكى تم وزن أحجام الرواسب (بالجرام) وتسجيل نسبتها المختلفة حسب مقياس وحدات فاى (ϕ)[1] ووضعت فى الجدول التالى:

(١) مقياس فاى (ϕ) هو وسيلة لتصنيف الرواسب حسب حجم حبيباتها. وهو مأخوذ عن نظام ونتوورث الملليمترى Wentworth's millimetre system الذى يقوم على أساس لوغاريتيمى. أى أنه عبارة عن التحويل اللوغاريتمى لحجم الرواسب بالملليمتر، أو هو يساوى اللوغاريتم السالب (للأساس ٢) لحجم الرواسب بالملليمتر (ϕ = − $_٢$لو ر، حيث ر = قطر الحبيبات بالملليمتر) .

جدول رقم (١٠ – ١٣): وزن حجم الحبيبات (بالجرام) لست عينات من الرواسب الشاطئية

الشاطئ	أ		ب		حـ		د		هـ		و	
وحدات الحجم	الوزن	الرتبة	الوزن	الرتبة	الوزن	الرتبة	الوزن	الرتبة	الوزن	الرتبة	الوزن	الرتبة
-٢ إلى -١	١٠٥	١٠,٥	٢٨٢	٢٩	١٢٥	١٦	٣٠٧	٣١	٠٠	٠٠	٢٩١	٣٠
-١ إلى صفر	١٢٣	١٥	٢٢١	٢٥	٠٠	٠٠	٨٥	٩	٦٠	٤	٣٢٤	٣٢
صفر إلى ١	٠	٠٠	٢٤٨	٢٦	٠٠	٠٠	٢٠١	٢٤	١٠٥	١٠,٥	١١٢	١٣
١ إلى ٢	١٤٥	١٨	٢٥٢	٢٧	٥٥	٢,٥	١٨٤	٢٣	٤٥	١	١٠٨	١٢
٢ إلى ٣	٧٨	٧	١٤٨	١٩	١٢١	١٤	٣٥٥	٣٣,٥	٣٥٥	٢٣,٥	١٦٥	٢١
٣ إلى ٤	١٦٠	٢٠	١٣١	١٧	٧٧	٦	١٧٥	٢٢	٠٠	٠٠	٢٨٠	٣٥
٤ إلى ٥	٨٢	٨	٢٧٣	٢٨	٦٧	٤			٥٥	٢,٥	٠٠	٠٠
المجموع	ر = ٧٨,٥ $ن_{١}$ = ٦		$ر_{٢}$ = ١٧١ $ن_{٢}$ = ٧		$ر_{٣}$ = ٤٣,٥ $ن_{٣}$ = ٥		$ر_{٤}$ = ١٤٢,٥ $ن_{٤}$ = ٥		$ر_{٥}$ = ٥١,٥ $ن_{٥}$ = ٦		$ر_{٦}$ = ١٤٣ $ن_{٦}$ = ٦	

١- الفروض: H_0 : لايوجد إختلاف بين أحجام الحبيبات لرواسب جميع العينات، أو بعبارة أخرى أن مقدار الاختلاف بين الأحجام المختلفة لحبيبات الرواسب الشاطئية يرجع إلى الصدفة وتبعاً لذلك فإن العينات الست تمثل مجتمعاً واحداً من الرواسب أو ستة مجتمعات لها نفس المتوسط لحجم حبيبات الرواسب.

H_1 : أن حجم حبيبات الرواسب يختلف من عينة لأخرى بما يؤكد أنها تمثل أكثر من مجتمع واحد من مجتمعات الرواسب الشاطئية.

٢- مستوى الدلالة (α) = ٠٥,

٣- يمكن وضع الاختبار على الصورة الآتية:

$$\text{هـ} = \frac{١٢}{\text{ن}\,(\text{ن} + ١)} \times \text{مجـ}_{\text{ن}} \frac{\text{ر}^{٢}}{\text{ن}} - ٣\,(\text{ن} + ١)$$

٤- حيث أن عدد العينات (ن) = ٦ وعدد المفردات (القياسات «ن») فى كل عينة لايقل عن ٥ مفردات فإن قيمة «هـ» لها توزيع احتمالى يتطابق مع توزيع مربع كاى بدرجات الحرية (ن - ١) أى ٦ - ١ = ٥.

٥- منطقة الرفض: على الأساس الموضوع سابقاً من (١) إلى (٤) نجد أن قيمة «هـ» الحرجة التى تحدد منطقة الرفض من جدول توزيع مربع كاى بدرجة الحرية ٤ ومستوى الدلالة ٠٥, هى ١١,٠٧ ونقبل فرض العدم فقط عندما تكون قيمة «هـ» المحسوبة من البيانات المشاهدة أقل من القيمة ١١,٠٧.

٦- حساب قيمة «هـ» من واقع البيانات المشاهدة: نتبع نفس الطريقة السابقة فى المثال (٢) لترتيب وجمع الرتب فنحصل على ما يلى:

$\text{ر}_{١}$ = ٧٨,٥	$\text{ر}_{٢}$ = ١٧١	$\text{ر}_{٣}$ = ٤٣,٥
$\text{ن}_{١}$ = ٦	$\text{ن}_{٢}$ = ٧	$\text{ن}_{٣}$ = ٥
$\text{ن}_{٤}$ = ١٤٢,٥	$\text{ر}_{٥}$ = ٥١,٥	$\text{ر}_{٦}$ = ١٤٣
$\text{ن}_{٤}$ = ٦	$\text{ن}_{٥}$ = ٥	$\text{ن}_{٦}$ = ٦

ن = ٣٥ ، ق = ٥

فتكون:

$$\text{هـ} = \frac{١٢}{١٨ \times ١٩} \times \left(\frac{(٧٨,٥)^٢}{٦} + \frac{(١٧١)^٢}{٧} + \frac{(٤٣,٥)^٢}{٥} + \frac{(١٤٢,٥)^٢}{٦} + \frac{(٥١,٥)^٢}{٥} + \frac{(١٤٣)^٢}{٦} \right) - ٣ \times ١٩$$

$$= (,٠٠٩٥ \times ١٢٩٠٥,٧٦٩) - ١٠٨$$

$$= ١٤,٦٠٥$$

٧- الاستنتاج: تحت شروط الاختبار نجد أن قيمة «هـ» المحسوبة أكبر من القيمة ١١,٠٧ بمستوى الدلالة ٠٥, ودرجة الحرية ٥ ولذلك يرفض فرض العدم القائل بأنه لايوجد فرق جوهرى بين حجم الرواسب فى العينات الست، أو أن الفرق بينها لايرجع إلى الصدفة المطلقة بل هو فرق معنوى يؤكد وجود اختلافات حقيقية بين مجتمعات الرواسب الشاطئية التى أخذت منها العينات. ولكن إذا أخذنا مستوى الدلالة ٠١, كأساس للاختبار فإن قيمة «هـ» الحرجة التى تحدد منطقة الرفض بنفس درجات الحرية هى ١٥,٠٩، وبالتالى تكون قيمة «هـ» المحسوبة أصغر من هذه القيمة مما يترتب عليه قبول فرض العدم السابق. ويعنى قبول فرض العدم فى هذه الحالة أن الفرق بين أحجام الرواسب فى العينات يرجع إلى الصدفة المطلقة والناتج من خطأ المعاينة. وهذا يعنى بشكل آخر بالنسبة للحالة الأولى بأن هناك احتمالا مقداره ٩٥٪ بأن لايكون الاختلاف بين أحجام الرواسب قد حدث بفعل الصدفة، أما بالنسبة للحالة الثانية، فإن الاختبار يعنى أننا غير متأكدين باحتمال مقداره ٩٩٪ أن هذا الاختلاف لم يحدث بفعل الصدفة. وفى مثل هذه الحالات يجب على الباحث أن يجمع بيانات أخرى عن الرواسب ولكن من عينة أكبر من العينة السابقة.

وكما لاحظنا فى المثالين (٢)، (٣) الأخيرين أن بكل منها عدد من الرتب

المتكافئة (المتساوية) Tied ranks، وفى مثل هذه الحالات لابد من تصحيح قيمة «هـ» المحسوبة حتى لاتتأثر النتائج المترتبة عليها. ويتم تصحيح قيمة «هـ» المحسوبة بقسمتها على عامل التصحيح التالى:

$$\text{عامل تصحيح قيمة هـ} = ١ - \frac{\text{مجـ}(ك^{٣} - ك)}{ن^{٣} - ن}$$

حيث ك هى عدد المفردات ذات الرتب المتكافئة بين مفردات كل عينة على حدة، ن هى العدد الكلى للمفردات فى كل العينات.

وبتطبيق عامل التصحيح السابق على كل من قيمتى «هـ» المحسوبتين فى المثالين (٢)، (٣) نجد أن:

أولاً: فى المثال رقم (٢) يوجد مفردتان فقط متساويتان فى الرتبة وبذلك فإن عامل التصحيح فى هذه الحالة يساوى:

$$١ - \frac{[(٢)^{٣} - ٢]}{(١٨)^{٣} - ١٨} = ١ - \frac{٦}{٥٨١٤} = ٠,٩٩٩$$

وتصبح قيمة «هـ» (٨,٩٠) المصححة:

$$\text{«هـ» المصححة} = \frac{٨,٩٠}{,٩٩٩} = ٨,٩١$$

ثانياً: فى المثال رقم (٣) يوجد ثلاث عينات بكل منها مفردتان متساويتان فى الرتبة وبذلك فإن عامل التصحيح فى هذه الحالة يساوى:

$$= ١ - \frac{[(٢^{٣} - ٢) + (٢^{٣} - ٢) + (٢^{٣} - ٢)]}{(٣٥)^{٣} - ٣٥}$$

$$= 1 - \frac{١٨}{٤٢٨٤٠} = ,٩٩٩٦$$

وتصبح قيمة «هـ» (١٤,٦٠٥) المصححة هي:

$$\text{«هـ» المصححة} = \frac{١٤,٦٠٥}{,٩٩٩٦} = ١٤,٦١١$$

ويلاحظ أن تأثير عامل التصحيح على قيمة «هـ» في المثالين صغيراً جداً ويكون ذلك صحيحاً حتى إذا وجد عدد كبير من المفردات المتكافئة في رتبتها بين المفردات الكلية للعينات. والغرض الأساسي من عامل التصحيح هو جعل قيمة «هـ» أكبر من القيمة المحسوبة لها مما يزيد من فرصة رفض فرض العدم. أو بعبارة أخرى إذا أهملنا عامل التصحيح السابق، وبصفة خاصة إذا كان ترتيب البيانات ينتج عنه أن ٢٥ ٪ أو أكثر من المفردات تأخذ رتباً متكافئة، فإن نتيجة اختبار «هـ» تصبح مضللة ويجب معالجتها إحصائياً بكل الحذر والحيطة، أما إذا قل كثيراً عدد الرتب المتساوية بين المفردات فإن تأثير عامل التصحيح يصبح بسيطاً جداً، كما لاحظنا، وبالتالي يمكن إهماله.



الباب الرابع
أساليب قياس العلاقات والتغيرات

مقدمة

الفصل الحادى عشر: تحليل الإرتباط

الفصل الثانى عشر: تحليل الإنحدار

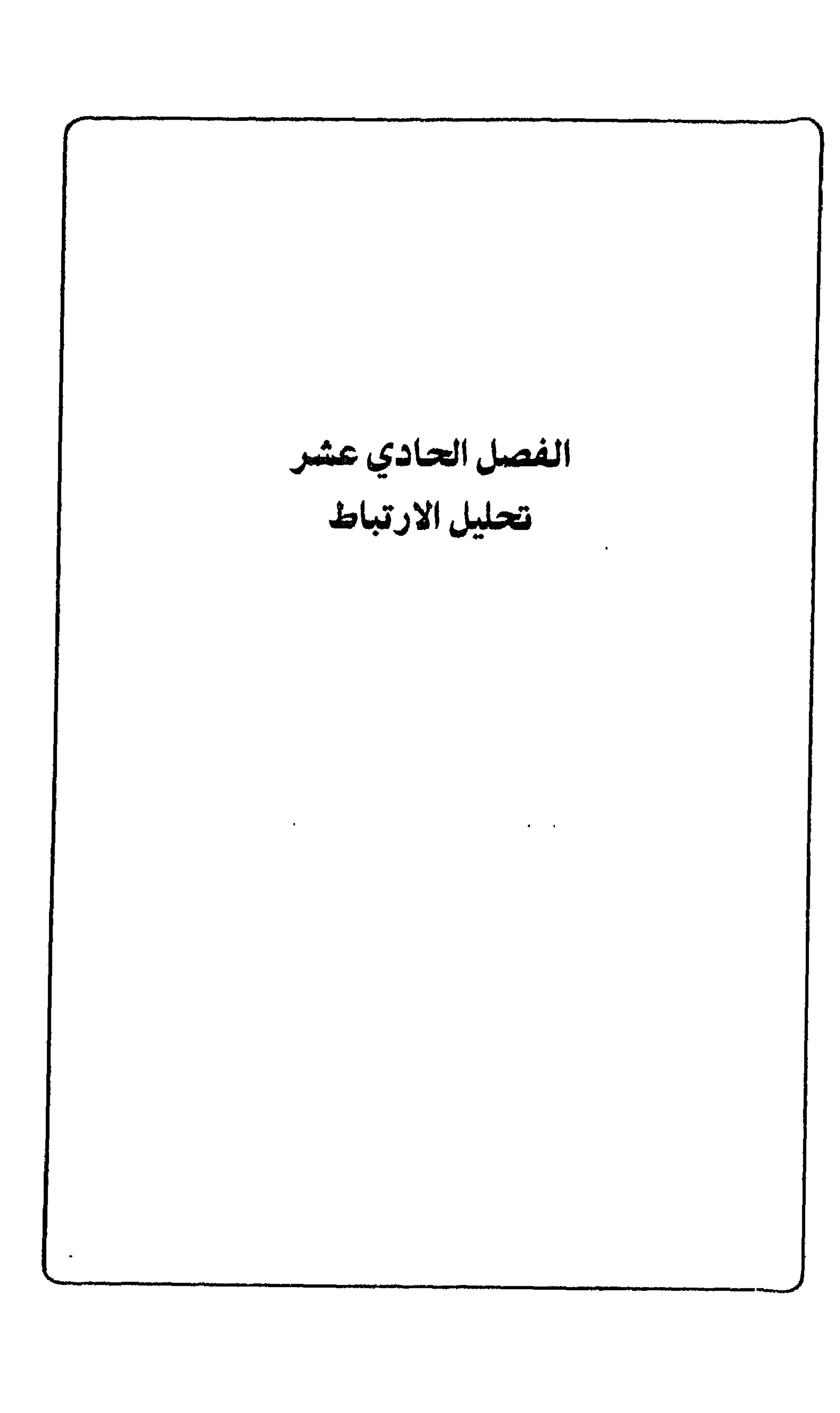
الفصل الثالث عشر: تحليل السلاسل الزمنية

Converted by Tiff Combine - (no stamps are applied by registered version)

مقدمة:

إن مهمة الباحث الجغرافي تنحصر في معالجة ووصف بيانات المتغيرات (الظواهر) الطبيعية والبشرية الموجودة على سطح الأرض لدراسة مواقعها ومقارنة توزيعاتها والعلاقات التي تربطها البعض. وقد شرحنا في فصول الأبواب لسابقة من هذا الكتاب الأساليب الكمية الخاصة بوصف وقياس وتحديد درجة الاختلاف بين عينتين أو أكثر على أساس تحديد معالم وخصائص متغير واحد (ظاهرة واحدة) فقط غير أن أسلوب التحليل الكمي لايقتصر على تحليل كل متغير بمفرده أو في صورة مستقلة عن المتغيرات الأخرى، بل يعطي لنا نظرية كاملة في مجال العلاقة بين متغيرين وقياس الترابط بينهما، ولقد كانت ومازالت مقارنة خرائط التوزيعات الجغرافية من أهم الوسائل لدراسة الارتباط والتلازم بين المتغيرات الجغرافية المختلفة. غير أن كثيرا من الدراسات والبحوث الجغرافية تظهر هذه الارتباطات على أنها علاقات سببية، مما كان له أكبر الأثر في توجيه اللوم للجغرافيين واتهامهم بأنهم كانوا يطوفون بأعينهم على الخرائط ليخرجوا منه باستنتاجات سريعة، أو كانوا يحلقون في سماء مكونات البيئة ونشاط الإنسان للتدليل على كثير من الحقائق أثبت التحليل الكمي الحديث عدم دقة وخطورة الإتماد عليها في وصف ودراسة العلاقات المختلفة والمتشابكة بين المتغيرات الجغرافية المتعددة.

وفي هذا الباب والذي يشتمل على ثلاثة فصول - سنحاول إلقاء الضوء على الأساليب الكمية التي تقيس درجة واتجاه العلاقة بين المتغيرات الجغرافية في الفصلين الأول والثاني. وتتمثل هذه الأساليب في تحليل الارتباط وتحليل الانحدار. أما في الفصل الأخير فستتعرض لدراسة بيانات الظاهرة الجغرافية على مدى فترات زمنية متتابعة (أو مايعرف بتحليل السلاسل الزمنية) بغرض وصف خط سير (أي تطور أو نمو) هذه الظاهرة في مدى زمني محدد. ثم نتعرف على الأساليب المتبعة في قياس التغيرات المختلفة التي قد تطرأ على تطور الظاهرة بفعل المؤثرات العديدة والمتباينة بهدف الاستفادة من هذه الأساليب في التنبؤ بما يمكن أن تكون عليه بيانات الظاهرة قيد التحليل سواء في الماضي أو في المستقبل.



الفصل الحادي عشر
تحليل الارتباط

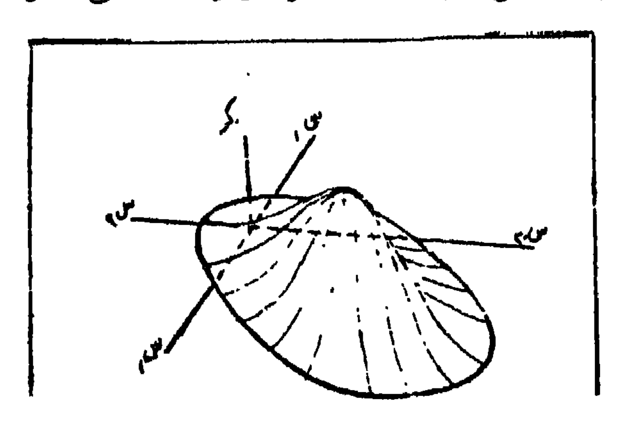


تحليل الارتباط

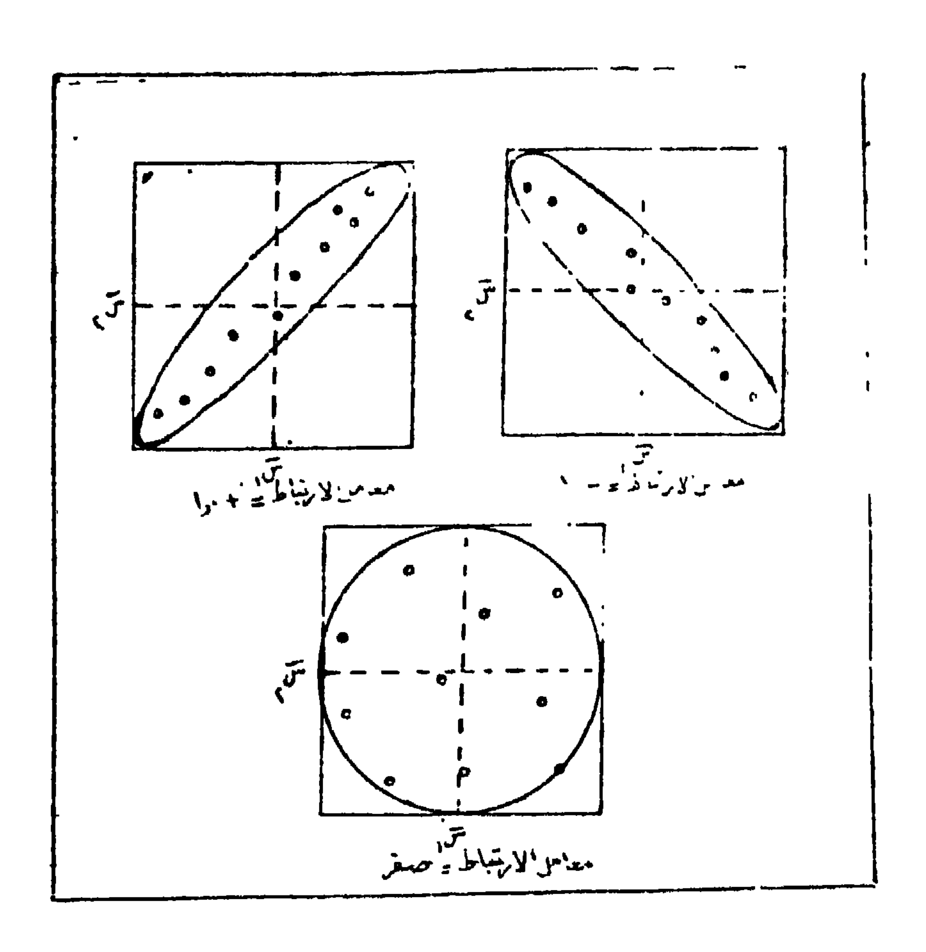
سبق القول بأن الأسلوب الكمى الحديث يساعد الباحث الجغرافى على الوصول إلى أهدافه العلمية بوسائل أكثر دقة من الأسلوب الوصفى التقليدى. ومن هنا نجد أن الجغرافيين الآن يهتمون بتطبيق أسلوباً كمياً معيناً - هو تحليل الارتباط - على بيانات الظواهر (المتغيرات) الجغرافية لكى يعرفوا به إن كان ثمة علاقة أو ارتباط بين ظاهرتين معينتين ولتحديد ما إذا كانت هذه العلاقة تعود إلى تلازم بين الظاهرتين أو إلى اختلافات ترجع إلى الصدفة المطلقة نتيجة خطأ المعاينة. ونقصد بتحليل الارتباط أنه الأسلوب الذى يقيس درجة الترابط بين ظاهرتين (متغيرين) إذا كانت العلاقة بينهما ليست علاقة دالية (أى أن التغير فى أحد الظاهرتين بالتلازم - لعلاقة بين طول النهر وعرضه فإن أى تغير فى أحدهما لا يسبب تغيراً فى الآخر ولكن المتغيرين مرتبطين بحجم الحوض النهرى وخصائصه الجيورمورفولوجية.

ويشترط عند تحليل الارتباط أن يتبع توزيع كل متغير من المتغيرين التوزيع المعتدل، أما توزيع المتغيرين معاً فإنه يشترط أن يتبع توزيع Bivariate Normal Distribution (شكل رقم ١١-١)، الذى يشبه الناقوس حيث تمثل قاعدته المتغيرين وارتفاعه يمثل التكرار أو لكلا المتغيرين، وذلك حتى يمكن تطبيق الأساليب لبارامترية الخاصة بقياس درجة الارتباط. بين المتغيرين، بينما إذا كان توزيع أحد المتغيرين، أو كلاهما، لا يتبع التوزيع المعتدل فإنه يمكن تطبيق أساليب أخرى غير بارامترية لقياس درجة الارتباط بين المتغيرين. ويلاحظ من الشكل رقم (١١-١) أن أى قطع عمودى على المستوى الذى يمثل المتغير الأول ($س_١$) ينتج عنه منحنى معتدل يمثل توزيعاً للمتغير الثانى ($س_٢$) والعكس صحيح. أما إذا كان القطع موازياً للمستويين $س_١$، $س_٢$ فإن ذلك يعطى شكل قطع ناقص Ellips يمكن اعتباره دليلاً على نوع وقوة الإرتباط (العلاقة) بين المتغيرين (شكل رقم ١١-١). فإذا

كان القطع الناتج يأخذ اتجاهاً معيناً فإن ذلك يدل على نوع الإرتباط فإما يكون الإرتباط موجب (أى علاقة طردية) ؛ أى أن تزايد قيم أحد المتغيرين يصحبه تزايد فى قيم المتغير الآخر والعكس - وهناك متغيرات كثيرة تتبع هذا النوع من الإرتباط نذكر منها على سبيل المثال لا الحصر تزايد الأمطار والإنتاج الزراعى فى المناطق الجافة وشبه الجافة؛ بمعنى أن أى تزايد فى الأمطار يصحبه تزايد فى إنتاج المحاصيل، شدة انحدار السفوح والتعرية؛ إذا أنه كلما زادت درجة الانحدار زادت شدة التعرية، زيادة سرعة المياه فى الأنهار وكمية الرواسب المحموله، والكفاءة الإنتاجية للعمال والإنتاج الصناعى؛ فكلما تحسنت الكفاءة الإنتاجية كلما زاد الإنتاج وكلما ضعفت الكفاءة قل الإنتاج. وإما يكون الإرتباط سالب (علاقة عكسية) ؛ أى تزايد قيم أحد المتغيرين يصحبه أنخفاض فى قيم المتغير الآخر. ومن أمثلة هذا النوع من العلاقة العكسية كثافة السكان والعبد عن وسط المدينة، كذلك أسعار الأراضى والبعد عن قلب المدينة التجارى - فالواضح أن تزايد المسافة بالبعد عن وسط المدينة وقلبها التجارى يرتبط به تغير عكسى فى كثافة السكان وأسعار الأراضى، أى أن الكثافة وأسعار الأراضى تقل كلما تزايد طول المسافة عن وسط المدينة وإما يكون الإرتباط معدوم؛ أى أنه ليس هناك علاقة أو ارتباط (موجب أو سالب) بين المتغيرين ون أمثلة ذلك البعد عن وسط المدينة وكمية الإنتاج الصناعى، أو البعد عن قلب المدينة التجارى وسرعة المياه فى المجارى النهرية.



شكل رقم (١١-١): التوزيع المعتدل للمتغيرين
س١، س٢ معا Biveriate Normal Distribution



شكل رقم (١١-٢): قوة الإرتباط بين المتغيرين $س_١$، $س_٢$
كما يوضحها شكل إنتشار المفردات لكل منهما

وكما يدل اتجاه القطع الناقص – الذى ينتج عن قطع توزيع المتغيرين معا قطعاً موازياً للسطحين $س_١$، $س_٢$ فى الشكل رقم (١١-١) على نوع الإرتباط أو العلاقة، يدل شكل القطع الناقص نفسه على قوة الإرتباط أو العلاقة. فإذا كان شكل القطع الناقص ضيق ومحدود، أى أن بيانات المتغيرين تقع فى مجال انتشار متقارب أو على امتداد خط مستقين، فهذا يدل على أن هناك علاقة قوية بين المتغيرين. أما إذا كان شكل القطع بعض الشئ بحيث تظهر البيانات متباعده تباعداً

طفيفاً ولكن حول خط مستقيم، دل ذلك على وجود علاقة ضعيفة. أما فى حالة إذا كان شكل القطع دائرى، أى أن هناك تباعداً كبيراً بين البيانات بحيث يتعذر وقوعها على امتداد خط مستقيم، فإن ذلك يدل عدم وجود علاقة بين المتغيرين أويدل على استقلال كل منهما عن الآخر.

ومما تجدر الإشارة إليه بشأن العلاقات الإرتباطية واتجاهها بين المتغيرات، أنه فى حالة إثبات وجود علاقة قوية أيا كان نوعها بين متغيرين، أو بعبارة أخرى أنه لو تعرفنا خلال الاختبار بأن العلاقة بين متغيرين لانعود إلى الصدفة فإن هذا لايعنى بأن هناك علاقة سببية بينهما، ذلك لأن وجود الإرتباط لايعنى بالضرورة وجود علاقة سببية (علاقة تبعية مباشرة) بين المتغيرين ، بل أن كل ما يكشف عنه (الإختبار) هو أن المتغيرين متلازمان تلازماً شديداً، مما يتيح الفرصة ويفسح المجال بعد ذلك البحث عن العلاقة الحقيقية الواقعة بينهما، ولكن إذا كان هناك علاقة سببية بين المتغيرين فلابد أن يكون هناك ترابط بينهما. فمثلاً عند حساب الإرتباط بين متوسط محصول القمح ودرجة الإصابة بدودة ورق القطن وجد أن هناك ارتباطاً قوياً بينهما، ولكن من الصعب تفسير هذه العلاقة من الناحية النطقية إذ لا يوجد سبب واحد بين هذين المتغيرين يؤدى إلى وجود هذه العلاقة، ولكن هذا الإرتباط (العلاقة) الزائف سببه الظروف المناخية الملائمة للنمو نبات القمح أثناء فصل الشتاء تؤثر على أعداد دودة ورق القطن. كذلك قد يكون هناك ارتباط قوى بين عدد الكتب المنشورة فى كل سنة وعدد مباريات كرة القدم فى كل سنة، مثل هذا الإرتباط يشار إليه بأنه ارتباط لامعنى له أو ارتباط زائف. ومن هنا يمكن القول أن الترابط ليس شرطاً للعلاقة السببية ولكن السببية شرط للترابط.

مقاييس الإرتباط Measures Of Corrclation

يمكن أن نحدد بصورة وصفية مدى جودة وصف العلاقة أو الإرتباط بين المتغيرات بملاحظة الشكل البيانى الذى يوضح مجال انتشار قيم المتغيرين فى نظام الاحداثيات المتعامدة، وهو الشكل الذى يعرف باسم «شكل الانتشار» (راجع لفصل الثالث الخاص بطرق العرض البيانى للبيانات الاحصائية). ولكن يمكن معرفة مدى جودة تعبير خط مستقيم عن العلاقة بين المتغيرات عن طريق حساب

معامل ارتباط Correlation Coefficient والذى وعلى أساسه يستخلص الباحث الجغرافى النتائج ويتخذ القرارات الخاصة بتوضيح العلاقات بين المتغيرات (الطبيعية أو البشرية) الجغرافية.

حساب معامل الإرتباط:

يستخدم مصطلح «معامل الإرتباط» ليعنى الإرتباط الخطى (أو العلاقة الخطية) بين متغيرين، وهو لذلك يستخدم - كمقياس إحصائى - لتحديد نوع العلاقة وقوتها بين المتغيرات. وتتراوح قيمة معامل الإرتباط المحسوبة بين (-١,٠٠)، (+١,٠٠)، حيث تشير القيمة (-١,٠٠) إلى وجود ارتباط سالب أو عكسى تام Negative (Inverse) Correlation، أما (+١) فترمز إلى وجود علاقة ارتباط طردية أو موجبة تامة. Positive (Direct) Correlation والقيمتان تدلان على أن جميع القيم الممثلة للعلاقة بين المتغيرين تقع على خط مستقيم. وكلما أخذت القيم تنحرف عن الخط المستقيم كلما قلت قيمة معامل الإرتباط عن القيمتين السابقتين بحكم ضعف العلاقة بين قيم المتغيرين، حتى إذا وصلت إلى درجة الصفر دل ذلك على انعدام العلاقة الإرتباطية. وإذا كانت الإشارات (+، -) تشير إلى نوع الإرتباط الخطى، فإن قيمة معامل الإرتباط لا تمييز لها أى أنها لاتعتمد على وحدات القياس المستخدمة. وينبغى أن نؤكدها هنا، مرة أخرى على أن معامل الإرتباط يقيس مدى جودة توفيق الصيغة الاحصائية المفترضة للبيانات، أو بعبارة أخرى أن قيمة معامل الإرتباط المحسوبة بين متغيرين تقيس فقط درجة العلاقة بينهما بالنسبة إلى نوع الصيغة الاحصائية المستخدمة.

وهناك صيغ مختلفة لحساب قيمة معامل الإرتباط (يرمز له بالرمز «ر») بين متغيرين، الا أن الصيغة التالية يفضلها كثير من الاحصائيين لأنها تعتمد على القيم الأصلية للمتغيرين (س١، س٢).

$$\text{ر} = \frac{\text{ن مجـ } \text{س}_1 \text{س}_2 - (\text{مجـ س}_1)(\text{مجـ س}_2)}{\sqrt{[\text{ن مجـ س}_1^{\,2} - (\text{مجـ س}_1)^2 \times \text{ن مجـ س}_2^{\,2} - (\text{مجـ س}_2)^2]}} \qquad (١١-١)$$

حيث ن هى عدد أزواج القيم للمتغيرين معاً.

وسنوضح كيفية حساب معامل الإرتباط باستخدام هذه الصيغة من المثال التالى:

مثال (١):

لدراسة العلاقة بين شكل الشاطئ ممثلاً بدرجات الانحدار، وشدة التعرية البحرية ممثلة بقوة الأمواج، أخذت عينة مكونة من ٦ قطاعات على الشاطئ وقيست درجة الانحدار وقوة الأمواج فكانت كمايلى:

قوة الأمواج (عشرات كجم م/م	١٠	١٥	١٨	٢٠	٢٤	٢٨
انحدار الشاطئ (بالدرجات)	٢	٣	٤	٤	٥	٦

والمطلوب حساب مقدار الإرتباط بين هاتين الظاهرتين:

لتسهيل عملية الحساب يتم ترتيب البيانات فى الجدول التالى:

جدول رقم (١١-١)

حساب الإرتباط بين قوة الأمواج وانحدار الشاطئ

قوة الأمواج	درجة الإنحدار			
$س_{١}$	$س_{٢}$	$س_{١}^{٢}$	$س_{٢}^{٢}$	$س_{١} س_{٢}$
١٠	٢	١٠٠	٤	٢٠
١٥	٣	٢٢٥	٩	٤٥
١٨	٤	٣٣٤	١٦	٧٢
٢٠	٤	٤٠٠	١٦	٨٠
٢٤	٥	٥٧٦	٢٥	١٢٠
٢٨	٦	٧٨٤	٢٦	١٦٨
١١٥	٢٤	٢٤٠٩	١٠٦	٥٠٥

معامل الإرتباط (ر) =

$$\frac{\text{ن مجـ } س_1 س_2 - (\text{مجـ } س_1)(\text{مجـ } س_2)}{\sqrt{[\text{ن مجـ } س_1^{٢} - (\text{مجـ } س_1)^{٢} \times \text{ن مجـ } س_2^{٢} - (\text{مجـ } س_2)^{٢}]}}$$

وحيث ن = ٦ فإن:

$$ر = \frac{٦ \times ٥٠٥ - (١١٥ \times ٢٤)}{\sqrt{[٦ \times ٢٤٠٩ - (١١٥)^{٢} \times (٦ \times ١٠٦) - (٢٤)^{٢}]}}$$

$$ر = \frac{٣٠٢٠ - ٢٧٦٠}{\sqrt{(١٤٤٥٤ - ١٣٢٢٥)(٦٣٦ - ٥٧٦)}}$$

$$ر = \frac{٢٧٠}{\sqrt{١٢٢٩ \times ٦٠}} = \frac{٢٧٠}{\sqrt{٧٣٧٤٠}}$$

$$ر = \frac{٢٧٠}{٢٧١,٦} = + ٠,٩٩٤١$$

أى أن معامل الإرتباط هو + ٠,٩٩٤١ قريب جداً من قيمة الإرتباط التام (+١) مما يدل على أن معامل الإرتباط قوى جداً. ويمكن أن نستنتج من ذلك أن هناك علاقة جغرافية وثيقة أو شديدة بين عنصرى قوة الأمواج وطبيعة شكل الشاطئ. ويلاحظ هنا أن العلاقة موجبة (علاقة طردية) أى أنه عندما تزداد قوة الأمواج يشتد انحدار الشاطئ.

مثال (٢):

فى دراسة لبيان العلاقة بين كثافة السكان والبعد عن قلب المدينة التجارى، قام باحث جغرافى بقياس المسافة من قلب المدينة (بالكيلومتر) وحساب كثافة السكان فى الكيلومتر المربع من قلب المدينة حتى خارجها فكانت البيانات التى حصل عليها كمايلى:

المسافة من قلب المدينة (كم)	١	٢	٣	٤	٥	٦	٧	٨
الكثافة (ألف شخص/ كم٢)	١٠	٦	٨	٦	٥	٣	٣	٢

ولحساب معامل الإرتباط بين هذين المتغيرين ترتب البيانات كما فى الجدول التالى:

جدول رقم (١١-٢)

حساب العلاقة بين كثافة السكان والمسافة من قلب المدينة

المسافة	الكثافة			
$س_{١}$	$س_{٢}$	$س_{١}^{٢}$	$س_{٢}^{٢}$	$س_{١} س_{٢}$
١	١٠	١	١٠٠	١٠
٢	٦	٤	٣٦	١٢
٣	٨	٩	٦٤	٢٤
٤	٦	١٦	٣٦	٢٤
٥	٥	٢٥	٢٥	٢٥
٦	٢	٣٦	٩	١٨
٧	٣	٤٩	٩	٢١
٨	٢	٦٤	٤	١٦
المجموع ٣٦	٤٣	٢٠٤	٢٨٣	١٥٠

وبتطبيق معادلة معامل الإرتباط السابقة (حيث ن = ٨) نحصل على قيمة المعامل وهى:

$$ر = \frac{٨ \times ١٥٠ - ٣٦ \times ٤٣}{\sqrt{[٨ \times ٢٠٤ - (٣٦)^{٢} \times (٨ \times ٢٨٣) - (٤٣)^{٢}]}}$$

$$= \sqrt{\frac{-٣٤٨}{٣٣٦ \times ٤١٥}} = \sqrt{\frac{-٣٤٨}{١٣٩٤٤٠}}$$

$$= \sqrt{\frac{-٢٤٨}{٣٧٣,٤٢}} = -٠,٩٣١٩$$

يتضح من نتيجة المعادلة أن هناك علاقة ارتباط قوية بين المسافة من قلب المدينة وكثافة السكان، ولكنها علاقة عكسية (سالبة)، أى أنه عندما تزداد المسافة من قلب المدينة تقل كثافة السكان فى الكيلو متر المربع تبعاً لذلك.

معامل ارتباط ضرب العزوم Product Moment Correlation Coefficient

يعد معامل ارتباط ضرب العزوم أو ما يعرف باسم معامل ارتباط بيرسون Pear-son orrelation Coefficient من أقوى الأساليب الإحصائية البارامترية (المعلمية) التى تقيس العلاقة بين متغيرين (ظاهرتين) يشترط فى بياناتها أن تكون من بيانات الفترة Intervally - Scaled.

وعلى الرغم من أن صيغة بيرسون لحساب معامل الإرتباط تعتبر كشفاً علمياً له أهمية كبيرة – لتحديد نوع ودرجة العلاقة بين المتغيرات – فى ميدان العلوم الطبيعية والبشرية، إلا أنها تحتاج إلى عمليات حسابية متعددة ودقيقة مما قد يؤدى إلى ارتكاب بعض الأخطاء التى قد تؤثر كثيراً على النتائج النهائية للدراسة. ولكن بفضل استخدام أجهزة الحاسبات الآلية Computers فقد أصبحت العمليات الحسابية لهذه الصيغة تتم بسهولة ويسر وبدون الوقوع فى أخطاء، مما أدى إلى أنها أصبحت تستخدم بكثرة فى البحوث الجغرافية.

وتعتمد أساساً صيغة بيرسون لمعامل الإرتباط للعينة على حساب إنحرافات قيم المتغيرات عن متوسطاتها الحسابية، وتكتب الصيغة بالشكل التالى:

$$ر_ب = \frac{\frac{١}{ن}\left[مجـ (س_١ - \bar{س}_٢) \times (س_٢ - \bar{س}_٢)\right]}{\sqrt{\frac{مجـ (س_١ - \bar{س}_٢)^٢}{ن} \times \frac{مجـ (س_٢ - \bar{س}_٢)^٢}{ن}}} \quad .... (١١-٢)$$

وحساب معامل الإرتباط بهذه الصيغة يكون صعباً، خاصة إذا كانت قيمة المتوسط الحسابي تحتوى على كسور عشرية مما قد يؤدى إلى تعقيد العمليات الحسابية وبالتالي يزيد من احتمالات الخطأ في النتيجة النهائية، لذلك فقد اشتقت عدة صيغ أخرى تكتب على النحو التالي:

$$ر_ب = \frac{\text{مجـ}\ س_1\ س_2 - ن\ \bar{س}_1\ \bar{س}_2}{\sqrt{(\text{مجـ}\ س_1 - ن\ \bar{س}_1)^2\ (\text{مجـ}\ س_2 - ن\ \bar{س}_2)^2}} \quad ......\ (11-3)$$

أو:

$$ر_ب = \frac{\frac{\text{مجـ}\ س_1\ س_2}{ن}}{\sqrt{\frac{\text{مجـ}(س_1)^2}{ن}} \times \sqrt{\frac{\text{مجـ}(س_2)^2}{ن} - \bar{س}^2}} \quad ......\ (11-4)$$

ويلاحظ أن حساب معامل الإرتباط من المعادلات السابقة يتطلب: حساب المتوسط الحسابي والانحراف المعياري للمتغيرين ($س_1$، $س_2$)، ومجموع حاصل ضرب كل من المتغيرين($س_1$، $س_2$).

وهناك صيغة أخرى لحساب معامل الإرتباط تعرف بالطريقة المختصرة التي تستفيد من الخصائص الحسابية لمعامل الإرتباط والتي يمكن أن تقلل إلى حد كبير من العمليات الحسابية والوقوع في الخطأ، هذه الخصائص هي: أن قيمة معامل الإرتباط لا تتغير إذا قسمنا أو ضربنا جميع قيم المتغير الأول على أو في عدد ثابت، أو إذا قسمنا أو ضربنا جيمع قيم المتغير الثاني على أو في أي عدد ثابت آخر. كما أن قيمة معامل الإرتباط لا تتغير إذا أضفنا أو طرحنا أي عدد ثابت إلى أو من جميع قيم المتغير الأول، أو إذا أضفنا أو طرحنا أي عدد ثابت آخر إلى أو من جميع قيم المتغير الثاني. وتكتب صيغة بيرسون المختصرة لتسهيل حساب معامل الإرتباط على النحو التالي:

$$\text{ر}_{\text{ب}} = \frac{\frac{1}{\text{ن}}\left(\text{مجـ}\ \text{ح}_{\text{س}1} \times \text{ح}_{\text{س}2}\right) - \bar{\text{ح}}_{\text{س}1} \times \bar{\text{ح}}_{\text{س}2}}{\sqrt{\frac{\text{مجـ}\ \text{ح}^2_{\text{س}1}}{\text{ن}} - \bar{\text{ح}}^2_{\text{س}1}} \times \sqrt{\frac{\text{مجـ}\ \text{ح}^2_{\text{س}2}}{\text{ن}} - \bar{\text{ح}}^2_{\text{س}2}}} \quad \ldots (11-5)$$

حيث ح هى انحراف قيم كل من المتغيرين عن المتوسط الحسابى لكل منهما، حَ هى متوسط الانحراف لقيم كل من المتغيرين عن المتوسط الحسابى لكل منهما، ن هى عدد أزواج المفردات للمتغيرين معاً.

ولحساب قيمة معامل الإرتباط (ر) بواسطة الصيغ السابقة نأخذ الأمثلة الآتية:

مثال (٢):

البيانات التالية تمثل مساحة حوض النهر وطوله لعينة من عشرة أنهار لأحد نظم التصريف النهرى فى منطقة ما. والمطلوب حساب معامل الإرتباط بين هذين المتغيرين.

طول النهر (كم) ٩,٤٠ ، ١٠,٨٥ ، ١٠,٦٩ ، ١٠,٩٧ ، ١١,٩٤ ، ١٨,٦٩ ، ١٧,٢٣ ، ١٨,٤١ ، ١٠,١٤ ، ١٠,٦٨.

مساحة حوض النهر (كم) ٤,٠٨ ، ٤,٥٣ ، ٤,٩٢ ، ٦,١٨ ، ٦,٢٢ ، ٦,٥٤ ، ٧,٥٥ ، ٨,٩٢ ، ١٠,١٦ ، ١٢,٥٠.

ولتوضيح خطوات حساب معامل ارتباط بيرسون بالصيغة (١١ -٢) يتم ترتيب المعلومات كما فى الجدول التالى:

جدول رقم (١١-٣)

طريقة حساب معامل الإرتباط بين طول النهر ($س_١$) ومساحة حوضه ($س_٢$)

$س_١$	$س_٢$	$س_١ - \bar{س}_١$	$س_٢ - \bar{س}_٢$	$(س_١ - \bar{س}_١)^٢$	$(س_٢ - \bar{س}_٢)^٢$	$(س_١ - \bar{س}_١) \times (س_٢ - \bar{س}_٢)$
٩,٤٠	٤,٠٨	٥,٥٠-	٣,٠٨-	٣٠,٢٥	٩,٤٩	١٦,٩٤
١٠,٨٥	٤,٥٣	٤,٠٥-	٢,٦٣-	١٦,٤٠	٦,٩٢	١٠,٦٥
١٠,٦٩	٤,٩٢	٤,٢١-	٢,٢٤-	١٧,٧٢	٥,٠٢	٩,٤٣
١٠,٩٧	٦,١٨	٣,٩٣-	٠,٩٨-	١٥,٤٤	٠,٩٦	٣,٨٥
١١,٩٤	٦,٢٢	٢,٩٦-	٠,٩٢-	٨,٧٦	٣,٦٩	٢,٧٢
١٨,٦٩	٦,٥٤	٣,٧٩	٠,٦٢-	١٤,٣٦	٠,٣٨	٢,٣٥
١٧,٢٣	٧,٥٥	٢,٣٣	٠,٧٩	٥,٣٤	٠,١٥	٠,٩٠
١٨,٤١	٨,٩٢	٣,٢٤	١,٧٦	١٠,٥٠	٣,١٠	٥,٧٠
٢٠,١٤	١٠,١٦	٥,٢٤	٣,٠٠	٢٧,٤٦	٩,٠٠	١٥,٧٢
٢٠,٦٨	١٢,٥٠	٥,٧٨	٥,٣٤	٣٣,٤١	٢٨,٥٢	٣٠,٨٧
١٤٩,٠٠	٧٠,٦٠				٦٧,٢٣	٩٩,١٣

$\bar{س}_١ = \frac{١٤٩}{١٠} = ١٤,٩$ ، $\bar{س}_٢ = \frac{١٤٩}{١٠} = ٧,١٦$

ومن البيانات السابقة يمكن حساب معامل ارتباط العينة (ر) بالمعادلة رقم (١١-٢) كمايلى:

$$= \frac{\frac{١}{١٠} \times ٩٩,١٣}{\sqrt{\frac{١٧٩,٦٤}{١٠}} \times \sqrt{\frac{٦٧,٢٣}{١٠}}}$$

$$= \frac{٩,٩١٣}{\sqrt{١٢٠,٧٧٢}} = \frac{٩,٩١٣}{١٠,٩٨٩} = +٠,٩٠٢$$

وحيث أن قيمة معامل الإرتباط (+ ٠,٩٠٢) قريبة ن العلاقة الإرتباطية التامة والتى توضح أن العلاقة بين طول النهر ومساحة حوضه هى علاقة موجبة طردية، بمعنى أن هذين المتغيرين يزدادان أو يتناقصان فى قيمتيهما معاً.

مثال (٣):

البيـانات التـاليـة توضح إنتـاج الفـدان من محـصـول القـمح (بالأردب) فى منطقتين زراعيتين $س_١$ ، $س_٢$ خلال عشرة سنوات متتالية هى:

السنة	١٩٧١	٧٢	٧٣	٧٤	٧٥	٧٦	٧٧	٧٨	٧٩	١٩٨٠	المتوسط
المنطقة $س_١$	٢٥	٢٦	٣٤	٢٥	٢٤	٢٨	٢٧	٢٩	٢٨	٢٩	٢,٧٥
المنطقة $س_٢$	٢١	١٧	٣٥	٢١	١٩	٢٢	٢٦	٢٢	٢٦	٢٦	٢٣,٥

يلاحظ على هده البيانات أن إنتاج الفدان فى المنطقة $س_١$ يزيد فى السنوات الخمس الأخيرة عن متوسط الإنتاج فى الفترة كلها، بينما فى المنطقة $س_٢$ يزيد إنتاج الفدان فى أربع سنوات فقط عن متـوسط الإنتـاج فى فتـرة العشـرة سنوات. ويمكن إيجاد قيمة معامل الإرتباط بين الإنتاج فى المنطقتين باستخدام الصيغة رقم (١١-٣) وذلك على النحو التالى:

جدول رقم (١١-٤)

خطوات حساب معامل الإرتباط بين إنتاج القمح فى المنطقتين $س_{١}$ ، $س_{٢}$

السنة	$س_{١}$	$س_{٢}$	$س_{١}^{٢}$	$س_{٢}^{٢}$	$س_{١}$ $س_{٢}$
١٩٧١	٢٥	٢١	٦٢٥	٤٤١	٥٢٥
١٩٧٢	٢٦	١٧	٦٧٦	٢٨٩	٤٤٢
١٩٧٣	٣٤	٣٥	١١٥٦	١٢٢٥	١١٩٠
١٩٧٤	٢٥	٢١	٦٢٥	٤٤١	٥٢٥
١٩٧٥	٢٤	١٩	٥٧٦	٣٦١	٤٥٦
١٩٧٦	٢٨	٢٢	٧٨٤	٤٨٤	٦١٦
١٩٧٧	٢٧	٢٦	٧٢٩	٦٧٦	٧٠٢
١٩٧٨	٢٩	٢٢	٨٤١	٤٨٤	٦٣٨
١٩٧٩	٢٨	٢٦	٧٨٤	٦٧٦	٧٢٨
١٩٨٠	٢٩	٢٦	٨٤١	٦٧٦	٧٥٤
			٧٦٣٧	٥٧٥٣	٦٥٧٦

$\bar{س}_{١} = ٢٧,٥$ ، $\bar{س}_{٢} = ٢٣,٥$

وتكون قيمة معامل الإرتباط (ر) للبيانات السابقة باستخدام المعادلة (١١-٣) هو:

$$\therefore ر = \frac{٦٥٧٦ - ١٠ \times ٢٧,٥ \times ٢٣,٥}{\sqrt{[٧٦٣٧ - ١٠ \times (٢٧,٥)^{٢}]\ [٥٧٥٣ - ١٠\ (٢٣٥)^{٢}]}}$$

$$= \frac{١١٣,٥}{\sqrt{٧٤,٥ \times ٢٢٠,٥}}$$

$$= \frac{١١٣,٥٠٠}{١٣١,٠٤٣} = + ٠,٨٦٦$$

وتدل قيمة معامل الإرتباط المحسوبة (+ ٠,٨٦٦) على وجود علاقة طردية قوية بين إنتاج القمح فى المنطقتين خلال فترة العشرة سنوات قيد الدراسة وتجدر الإشارة هنا إلى أن قيمة معامل الإرتباط الذى حصلنا عليها لا توضح السبب فى وجود هذه العلاقة وسوف نتناول بعد قليل كيفية توضيح العلاقات الإحصائية المعنوية بين المتغيرات.

مثال (٤):

فى أحد مراكز البحوث الزراعية التى تهتم بتأثير ارتفاع سطح الأرض عن مستوى سطح البحر على الإنتاج الزراعى سجلت البيانات التالية عن إنتاج الفدان من القمح (بالأردب) وارتفاع الأرض المنتجة (بالأقدام):

الارتفاع بالقدم (س$_1$)	١٠٠	٢٠٠	٥٠٠	٧٠٠	٨٠٠	١٠٠٠	١٤٠٠	١٥٠٠	١٨٠٠	٢٠٠٠
إنتاج الفدان بالأردب (س$_2$)	٣٠	٣٠	٣١	٢٤	٢٦	٢٣	١٣	١٧	١٤	١٢

والمطلوب حساب معامل الإرتباط بين هذين المتغيرين. ولإجراء ذلك ترتب البيانات كما فى الجدول التالى:

جدول رقم (١١-٥)

حساب معامل الإرتباط بين ارتفاع سطح لأرض وانتاج القمح

الارتفاع بالقدم ($س_١$)	إنتاج الفدان بالأرداب $س_٢$	$س_١^٢$	$س_٢^٢$	$س_١ س_٢$
١٠٠	٣٠	١٠٠٠٠	٩٠٠	٣٠٠٠
٢٠٠	٣٠	١٠٠٠٠	٩٠٠	٦٠٠٠
٥٠٠	٣١	٢٥٠٠٠٠	٩٦١	١٥٥٠٠
٧٠٠	٢٤	٤٩٠٠٠٠	٥٧٦	١٦٨٠٠
٨٠٠	٢٦	٦٤٠٠٠٠	٦٧٦	٢٠٨٠٠
١٠٠٠	٢٣	١٠٠٠٠٠٠	٥٢٩	٢٣٠٠٠
١٤٠٠	١٣	١٩٦٠٠٠٠	١٦٩	١٨٢٠٠
١٥٠٠	١٧	٢٢٥٠٠٠٠	٢٨٩	٢٥٥٠٠
١٨٠٠	١٤	٣٢٤٠٠٠٠	١٩٦	٢٥٢٠٠
٢٠٠٠	١٢	٤٠٠٠٠٠٠	١٤٤	٢٤٠٠٠
المجموع ١٠٠٠٠	٢٢٠	١٣٨٨٠٠٠٠	٥٣٤٠	١٧٨٠٠٠

$$\bar{س}_١ = \frac{١٠٠٠٠}{١٠} = ١٠٠٠ \quad ، \quad \bar{س}_٢ = \frac{٢٢٠}{١٠} = ٢٢$$

وبإستخدام المعادلة رقم (١١-٤) يمكن الحصول على قيمة معامل الإرتباط كمايلى:

$$ر = \frac{\frac{١٧٨٠٠٠}{١٠} - ١٠٠٠ \times ٢٢}{\sqrt{\frac{١٣٨٨٠٠٠٠}{١٠} - (١٠٠٠)^٢} \times \sqrt{\frac{٥٣٤٠}{١٠} - (٢٢)^٢}}$$

$$\therefore = \frac{١٧٨٠٠ - ٢٢٠٠٠٠}{\sqrt{١٣٨٨٠٠٠ - ١٠٠٠٠٠٠} \times \sqrt{٥٣٤ - ٤٨٤}}$$

$$= \frac{- ٧٤,٥٤٢٠٠}{\sqrt{٦٢٢,٨٩ \times ٧٠٠٧}}$$

$$= \frac{- ٤٢٠٠}{٤٤٠٣,٨٣} = - ٠,٩٥٤$$

وتدل قيمة معامل الإرتباط (−٠,٩٥٤) على أن هناك علاقة عكسية قوية بين ارتفاع سطح الأرض عند مستوى سطح البحر وإنتاج الفدان من محصول القمح.

مثال (٥):

البيانات التالية توضح كمية الأمطار الساقطة (بالبوصات) وشدة سريان المياه (بالبوصات) فى أحد الأنهار بإحدى المناطق ذات المناخ الموسمى لفترة ست عشرة سنة :

كمية الأمطار (بالبوصات): ٤٦,٤، ٦٣,٠، ٤٨,٨، ٦٠,١، ٥٠,٦، ٥٧,٥، ٥٥,٥، ٥٧,٠، ٦٠,٨، ٤٨,٣، ٥٩,٠، ٤١,٠، ٦٦,٧، ٥٦,٤، ٥٨,٣، ٥٥,٧.

شدة سريان مياه النهر (بالبوصات): ٣١,٩، ٤٦,٨، ٣٤,٢، ٤٧,٥، ٣٥,٢، ٤٠,٥، ٤١,٣، ٤٣,٥، ٤٤,٨، ٣٨,٥، ٣٩,١، ٢٩,٥، ٤٦,٥، ٤٣,٤، ٤٠,٩، ٤١,٣.

ولإجراء حساب معامل الإرتباط بين هذين المتغيرين، نختار وسطاً فرضياً لكمية المطر ($س_1$) وليكن ٥٥ بوصة، ووسطاً فرضياً لشدة سريان مياه النهر وليكن ٤٠ بوصة. وفى مثل هذه الحالة نستخدم الطريقة المختصرة ولتسهيل العمليات الحسابية نتخلص من الكسور العشرية وذلك بالضرب فى ١٠ وذلك على النحو التالى:

جدول رقم (١١-٦)

خطوات حساب معامل الإرتباط بين كمية الأمطار وشدة سريان مياه النهر

ح س١ ١٠ (س١ − ٥٥)	ح س٢ ١٠ (س٢ − ٤٠)	ح س١ × ح س١	ح² س١	ح² س٢
− ٨٦	− ٨١	٦٩٦٦	٧٣٩٦	٦٥٦١
+ ٨٠	+ ٦٨	٥٤٤٠	٦٤٠٠	٤٦٢٤
− ٦٢	− ٥٨	٣٥٩٦	٢٨٤٤	٣٣٦٤
+ ٥١	+ ٧٥	٣٨٢٥	٢٦٠١	٥٦٢٥
− ٤٤	− ٤٨	٢١١٢	١٩٣٦	٢٣٠٤
+ ٢٥	+ ٥	١٢٥	٦٢٥	٢٥
+ ٥	+ ١٣	٦٥	٢٥	١٦٩
+ ٢٠	+ ٣٥	٧٠٠	٤٠٠	١٢٢٥
+ ٥٨	+ ٤٨	٢٧٨٤	٢٣٦٤	٢٣٠٤
− ٦٧	− ١٥	١٠٠٥	٤٤٨٩	٢٢٥
+ ٤٠	− ٩	٣٦٠	١٦٠٠	٨١
− ١٤٠	− ١٣٥	١٨٩٠٠	١٩٦٠٠	١٨٢٢٥
+ ١١٧	− ٦٥	٧٦٠٥	١٣٦٨٩	٤٢٢٥
+ ١٤	+ ٣٤	٤٧٦	١٩٦	١١٥٦
+ ٣٣	+ ٩	٢٩٧	١٠٨٩	٨١
+ ٧	+ ١٣	٩١	٤٩	١٦٩
+ ٥١ مجـ ح س١	+ ١٩ مجـ ح س٢	٥٣٦٢٧ مجـ ح س١ ح س٢	٦٧٣٠٢ مجـ ح² س١	٥٠٣٦٣ مجـ ح² س٢

وبتطبيق المعادلة رقم (١١–٥) يمكن حساب معامل الإرتباط (ر) بعد حساب حدود هذه المعادلة كمايلى:

$$\bar{ح}_{س_1} = \frac{مجـ\ ح\ س_1}{ن} = \frac{٥١}{١٦} = ٣,١٩$$

$$\bar{ح}_{س_2} = \frac{مجـ\ ح\ س_2}{ن} = \frac{١٩}{١٦} = ١,١٩$$

$$\bar{ح}_{س_1} \times \bar{ح}_{س_2} = ٣,١٩ \times ١,١٩ = ٣,٧٩$$

$$\frac{مجـ\ ح\ س_1\ ح\ س_2}{ن} = \frac{٥٣٦٢٧}{١٦} = ٣٣٥١,٦٩$$

$$\sqrt{\frac{مجـ\ ح^2\ س_1}{ن} - \bar{ح}^2_{س_1}} = \sqrt{\frac{٦٧٣٠٢}{١٦} - (٣,١٩)^2}$$

$$= ٦٤,٧٨$$

$$\sqrt{\frac{مجـ\ ح^2\ س_2}{ن} - \bar{ح}^2_{س_2}} = \sqrt{\frac{٥٠٣٦٣}{١٦} - (٣,١٩)^2}$$

$$= ٥٦,٠٩$$

$$\therefore ر = \frac{٣٣٥١,٦٩ - ٣,٧٩}{٦٤,٧٨ \times ٥٦,٠٩} = \frac{٣٣٤٧,٩٠}{٣٦٣٣,٥١} = +٠,٩٢١$$

أى أن هناك ارتباط طردى قوى بين كمية الأمطار الساقطة وشدة انسياب المياه فى النهر.

معامل ارتباط الرتب Rank Correlation Ceofficient

فى حالة إذا كان توزيع المفردات لمتغيرين أو لأحدهما فى عينة غير معتدل، أو

عندما لا يكون مثل هذا التحديد متاح، فبدلاً من استخدام القيم الخاصة بهما فإنه يمكن ترتيب أو حجمها أو غير ذلك. وفى مثل هذه الحالة يستخدم مقياساً آخر لحساب نوع ودرجة الترابط بين المتغيرين يعرف باسم «معامل ارتباط الرتب». وهو أداة إحصائية غير بارامترية (غير معلمية) تقيس العلاقة بين نوعين من البيانات الترتيبية لمتغيرين. كما يعد معامل ارتباط الرتب الأداة البديلة لمعامل ارتباط ضرب العزوم. وذلك في حالة عدم افتراض التوزيع المعتدل لبيانات كل من المتغيرين. أو بعبارة أخرى لايشترط فى حساب معامل ارتباط الرتب أن يكون توزيع البيانات الأصلة للمتغيرين معتدلاً، ولكنه يشترط أن تكون بيانات المتغيرين من نوع بيانات الرتب.

ومن الطبيعى أن يكون هناك اختلاف بين قيمتى معامل الإرتباط للقيم الأصلية (معامل ارتباط ضرب العزوم أو معامل ارتباط بيرسون) ومعامل الإرتباط للرتب. والسبب فى ذلك يرجع إلى استبدال القيم الأصلية للمفردات برتب خاصة، وفى هذه العملية بعض التقريب. كما أن معامل ارتباط الرتب أقل دقة من معامل ارتباط ضروب العزوم الذى يتأثر بأى تغير فى القيم الأصلية التى تسجل عن مفردات العينة، بينما لا يتأثر معامل ارتباط الرتب بذلك لأن زيادة قيمة مفردة أو نقصها ولو بسيطاً لا يغير من وضع المفردة (ترتيبها) داخل العينة.

معامل ارتباط سبيرمان Spearman's Rank Correlation Coefficient

يشترط عند حساب معامل الإرتباط بين متغيرين بطريقة سبيرمان أن لايقل عدد المفردات (الحالات) المكونة للعينة عن عشرة مفردات. وتعتمد طريقة حساب معامل ارتباط سبيرمان على إعطاء المفردات رتباً لتحل محل القيم العددية الأصلية، حسب الأهمية أو الحجم، لكل متغير من المتغيرين قيد التحليل. ويلزم حساب هذا المعامل أن نرتب القيم الزصلية للمفردات ترتيباً تصاعدياً أو تنازلياً، ثم نحسب الفروق بين الرتب لكل حالة من المتغيرين، ثم تربع هذه الفروق حتى نتخلص من الإشارات الحسابية. وبعد ذلك يمكن إيجاد قيمة معامل سبيرمان باستخدام الصيغة التالية:

$$ر_س = ١ - \frac{٦ \text{ مجـ ف}^{٢}}{ن^{٣} - ن} \quad \text{.....................} \quad (١١ - ٦)$$

أو:

$$ر_س = ١ - \frac{٦ \text{ مجـ ف}^{٢}}{ن (ن^{٢} - ١)} \quad \text{.....................} \quad (١١ - ٧)$$

حيث (ف) هى لفروق بين رتبتى كل حالة، ن هى هدد أزواج الرتب.

مثال (٦):

نفرض أن هناك عشر مناطق صناعية رتبت ترتيباً تصاعدياً أو تنازلياً من ناحية إنتاج الآلات الميكانيكية وإنتاج السيارات وكانت الرتب على النحو التالى:

المنطقة	أ	ب	جـ	د	هـ	و	ز	ح	ط	ك
الرتبة فى إنتاج الآلات الميكانيكية	٢	١	٤	٣	٥	٧	٦	٩	١٠	٨
الرتبة فى إنتاج السيارات	٢	٣	٥	١	٤	٩	٦	٧	٨	١٠

يلاحظ من البيانات أن المنطقة (ب) هى الأولى فى إنتاج الآلات الميكانيكية وثالثة فى إنتاج السيارات، بينما المنطقة (د) رتبتها الأولى فى إنتاج السيارات والثالثة فى إنتاج الآلات الميكانيكية.

ويمكن الحصول على قيمة معامل ارتباط سبيرمان الرتب باستخدام القانون السابق كمايلى:

جدول رقم (١١-٧)

خطوات حساب معامل ارتباط سبيرمان بين إنتاج
الآلات الميكانيكية وإنتاج السيارات

النقطة	رتبة إنتاج الآلات الميكانيكية	رتبة إنتاج السيارات	الفرقة (ف)	ف ٢
أ	٢	٢	صفر	صفر
ب	١	٣	- ٢	٤
جـ	٤	٥	- ١	١
د	٣	١	+ ٢	٤
هـ	٥	٤	+ ١	١
و	٧	٩	- ٢	٤
ز	٦	٦	صفر	صفر
ح	٩	٧	+ ٢	٤
ط	١٠	٨	+ ٢	٤
ك	٨	١٠	- ٢	٤
المجموع				٢٦

$$ر_س = ١ - \frac{٦ \times ٢٦}{١٠ (١٠٠ - ١)} = ١ - \frac{١٥٦}{٩٩٠}$$

$$= ١ - ٠,١٥٧ = +٠,٨٤٣$$

وتشير قيمة معامل الإرتباط (+ ٠,٨٤٣) فى هذه الحالة إلى وجود علاقة طردية قوية بين هذين المتغيرين.

مثال (٧):

نعود إلى المثال رقم (٢) الذى توضح بياناته مساحة حوض النهر وطوله لعينة من عشرة أنهار لتحسب منه معامل سبيرمان لارتباط الرتب:

جدول رقم (١١-٨)

حساب معامل ارتباط سبيرمان بين طول النهر ومساحة حوضه

(س١) الطول الكلى النهر(كم)	(س٢) مساحة حوض النهر (كم٢)	رتب س١	رتب س٢		ف٢
٩,٤٠	٤,٠٨	١ (أقصر نهر)	١ (أصغر مساحة)	صفر	صفر
١٠,٨٥	٤,٥٣	٣	٢	١-	١
١٠,٦٩	٤,٩٢	٢	٣	١+	١
١٠,٩٧	٦,١٨	٤	٤	صفر	صفر
١١,٩٤	٦,٢٢	٥	٥	صفر	صفر
١٨,٦٩	٦,٥٤	٨	٦	٢-	٤
١٧,٢٣	٧,٥٥	٦	٧	١+	١
١٨,٤١	٨,٩٢	٧	٨	١+	١
٢٠,١٤	١٠,١٦	٩	٩	صفر	صفر
٢٠,٦٨	١٢,٥٠	١٠ (أطول نهر)	١٠ (أكبر مساحة)	صفر	صفر
المجموع					٨

ومن تطبيق قانون سبيرمان:

$$ر_س = ١ - \frac{٦ \times ٨}{١٠ (١٠٠ - ١)}$$

$$= ١ - \frac{٤٨}{٩٩٠} = + ٠,٩٥١٥$$

أى أن قيمة معامل الإرتباط بين المتغيرين هى + ٠,٩٥١٥ وهى تدل على أن العلاقة بين المتغيرين علاقة طردية قوية، كما أنها قيمة تقترب كثيراً من قيمة معامل إرتباط بيرسون (ر = ٠,٩٠٢) السابق حسابها لنفس البيانات.

مثال (٨):

البيانات الآتية تمثل معدلات النمو السكانى والنسبة المئوية لما يخص الفرد من الإنتاج القومى فى ١٤ دولة من دول العالم فى عام ١٩٧٠ والمطلوب حساب نوع ودرجة الإرتباط بينهما باستخدام طريقة سبيرمان لارتباط الرتب.

جدول رقم (١١-٩)

حساب معامل ارتباط سبيرمان بين معدلات النمو السكاني والنسبة المئوية لما يخص الفرد من الإنتاج القومي في ١٤ دولة من دول العالم (عام ١٩٧٠)

الدولة	معدل النمو السكاني	الرتبة	النسبة المئوية من الإنتاج القومي للفرد	الرتبة	الفرقة (ف)	ف٢
البرازيل	٣,٠	٢	١,٦	١٠	-٨	٦٤,٠٠
نيجيريا	٢,٤	٤	-٠,٣	١٢,٥	-٨,٥	٧٢,٢٥
ألمانيا الغربية	١,٠	١١	٣,٤	٥,٥	+٥,٥	٣٠,٢٥
المملكة المتحدة	٠,٧	١٤	٢,٠	٨	+٦	٣٦,٠٠
إيطاليا	٠,٨	١٣	٤,٠	٣	+١٠	٥١٠٠,٠٠
فرنسا	١,١	٩,٥	٣,٧	٤	+٥,٥	٣٠,٢
المكسيك	٣,٥	١	٣,٤	٥,٥	-٤,٥	٢٠,٢٥
أسبانيا	٠,٩	١٢	٦,٥	١	+١١	١٢,٠٠
مصر	٢,٥	٣	١,١	١٠	-٧	٤٩,٠٠
بورما	٢,١	٦	١,٦	١٠	-٤	١٦,٠٠
يوغوسلافيا	١,١	٩,٥	٤,٢	٢	+٧,٥	٥٦,٢٥
أفغانستان	٢,٠	٧	-٠,٣	١٢,٥	-٥,٥	٣٠,٢٥
هولندا	١,٣	٨	٣,٠	٧	+١	١,٠٠
الجزائر	٢,٣	٥	-٣,٥	١٤	-٩	٨١,٠٠
المجموع						٧٠٧,٥٠٠

(المصدر: World Ban; Atlas, 1970)

$$ر_س = ١ - \frac{٦ \times ٧٠٧,٥}{١٤ (١٩٦ - ١)}$$

$$= ١ - \frac{٤٢٤٥}{٢٧٣٠} = ١ - ١,٥٥٥$$

$$= - ٠,٥٥٥$$

وتدل قيمة معامل الإرتباط المحسوبة (-٠,٥٥٥) على وجود علاقة عكسية بين المتغيرين، بمعنى أنه فى حالة زيادة النسبة المئوية لنمو السكان تقل النسبة المئوية لما يخص الفرد من الإنتاج القومي في هذه الدول. وإذا فحصنا البيانات فى الجدول سنرى أن الأرقام التى تدل على ارتفاع معدل النمو السكانى وانخفاض نسبة مايخص الفرد من لاإنتاج القومى تختص بها الدول النامية (مثل المكسك والبرازيل)، يعكس ماهو ملاحظ على نفس المعدلات بالنسبة للدول المتقدمة (مثل المملكة المتحدة، وألمانيا الغربية).

وعلى الرغم من أن صيغة سبيرمان ليست بدقة صيغة معامل ارتباط العزوم، إلا أنها بسيطة فى الاستخدام ويفضلها الكثير من الباحثين لأنها سهلة ولاتتطلب عمليات حسابية معقدة. ولكن هناك بعض العيوب التى تؤخذ على معامل سبيرمان لارتباط الرتب منها ماهو خاص بانعدام المعنى الطبيعى للفرق بين رتبتين، ومعنى تربيع هذا الفرق، ومنها مايتصل بالتتوزيع الذى نحصل عليه من العينات المختلفة لحساب قيمة هذا المعامل.

معامل كندال لارتباط الرتب Kendall's Rank Correlation Coefficient

يفضل استخدام معامل كندال كثيراً عن معامل سبيرمان فى قياس درجة الإرتباط لأنه أسهل فى حسابه، كما يمكن استخدامه فى حالة العينات الصغيرة (عدد المفردات أقل من ١٠). ويتطلب حساب هذا المعامل إعادة ترتيب رتب مفردات أحد المتغيرين حسب الترتيب لاعادى إما تصاعديا أو تنازلياً مع ترك رتب

المتغير الآخر بدون إعادة ترتيبها، ثم إيجاد قيمة معامل الإرتباط بينهما بعد تطبيق معادلة كندال الآتية:

$$\text{ر}_{\text{ك}} = \frac{\text{مجـ د}}{{}^{\text{١}}/_{\text{٢}}\,\text{ن}\,(\text{ن} - \text{١})} \quad \text{أو} \quad \frac{\text{٢ مجـ د}}{\text{ن}\,(\text{ن} - \text{١})} \quad \ldots\ldots\ldots \quad (\text{١٤}-\text{٨})$$

حيث د هى مجموع الفروق بين الرتب. ويمكن الحصول على أكبر عدد من فروق الرتب إذا كانت كل الرتب فى ترتيبها العادى، فإذا كان لدينا ن من الرتب فإن أكبر عدد ممكن من فروق الرتب هو ${}^{\text{١}}/_{\text{٢}}$ن (ن-١).

ويشبه معامل كندال لارتباط الرتب معامل سبيرمان السابق شرحه، من حيث أن قيمته تنحصر بين + ١,٠ ، - ١,٠ . ونحصل على القيمة ±١ إذا كانت الرتب المتناظرة متفقة تماماً ويسمى الإرتباط فى هذه الحالة بالإرتباط التام الموجب. أما إذا كانت الرتب عكسية تماماً فإننا نحصل على قيمة لمعامل الإرتباط تساوى -١,٠ ، ويسمى الإرتباط فى هذه الحالة بالإرتباط التام السالب.

وهناك طريقتان يمكن بواسطتهما حساب معامل كندال لارتباط الرتب، ونعرض فيمايلى خطوات الحساب التى تتبع فى كل منهما.

الطريقة الأولى:

ينحصر حساب معامل كندال بهذه الطريقة فى الخطوات الآتية التى نطبقها على المثال التالى:

مثال (٩):

لدراسة العلاقة الإرتباطية بين كمية المطر والارتفاع عن سطح البحر عن طريق استخدام معامل كندال، أخذت البيانات الموضحة فى الجدول التالى (جدول رقم ١٤-١٠) لستة من محطات الأرصاد الجوية فى احدى المناطق ذات الطبيعة الجبلية.

جدول رقم (١١-١٠)

طريقة حساب معامل كندال بين متغيرى المطر وارتفاع الأرض عن سطح البحر لستة محطات جوية

المحطة	الارتفاع ($س_1$) (متر)	كمية المطر ($س_2$) (سنتيمتر)	ترتيب ($س_1$)	ترتيب ($س_2$)
أ	٢٥٠٠	٢٠٠	١	٢
ب	١٨٠٠	٢٢٥	٢	١
جـ	٩٠٠	١٩٠	٣	٤
د	٦٠٠	٢١٠	٤	٢
هـ	٤٠٠	١٢٠	٥	٦
و	١٥٠	٨٥	٦	٦

١- ترتب المحطات ترتيباً عادياً تنازلياً (أو تصاعدياً) لمتغير واحد وهو فى مثالنا، الارتفاع ($س_1$).

٢- نضع ترتيب المحطات بالنسبة للمتغير الثانى (كمية المطر $س_2$) كما هى (أى بدون ترتيبها) مقابل ترتيب ارتفاعها، وذلك على النحو التالى:

المتغير ($س_1$)	١	٢	٣	٤	٥	٦
المتغير ($س_2$)	٢	١	٤	٣	٦	٦

٣- نبدأ بملاحظة ترتيب ($س_2$) غير المرتب ترتيباً عادياً بادئين من ترتيب البداية وهو (٢)، ثم نسجل لكل ترتيب اختلافه عن الترتيب الذى يليه بإعطاء القيمة + ١ للترتيب الأكبر، والقيمة -١ للترتيب الأصغر. وفى هذه الحالة نجد أن الرتب التى تزيد عن الرتبة ٢ هى ٣، ٤، ٥، ٦ والتى تقل عنها هى ١ وتكون القيم المقابلة لها مجتمعه هى (-١، + ١، + ١، + ١، + ١)

والمجموع الجبرى لهما يساوى +٣ (أى (+٤) + (−١) = +٣)، ولا يؤخذ هذا الترتيب فى الاعتبار مرة أخرى.

٤− وهكذا بالنسبة لباقى الرتب يمكن إيجاد القيمة المقابلة للرتب الأكبر والأصغر من الرتبة التالية للمتغير الثانى ($س_٢$). وفى النهاية تقوم بجمع القيم المقابلة لجميع رتب هذا المتغير، فنحصل على المجموع الكلى لقيم الاختلاف، والتى يرمز لها بالرمز (د)، وذلك على النحو التالى:

ترتيب ($س_٢$)	القيم المقابلة	
٢	(+٤) + (−١)	= +٣
١	(٤)	= +٤
٤	(−١) + (+٢)	= +١
٣	(+٢)	= +٢
٥	(+١)	= +١
٦	(صفر)	= صفر
المجموع (د)		= +١١

٥− وبما أن معامل كندال لارتباط الرتب عبارة عن مقياس لاختلاف رتب أو لفروق ترتيب أحد المتغيرين عندما يكون المتغير الآخر مرتباً ترتيباً (أى فى تسلسل منتظم)، فإن قيمة المعامل يمكـن الحصول عليها بقسمة المجموع العام للقيم المحسوب من اختلاف الرتب (د) على المجموع الأكبر المحتمل $\frac{١}{٢}$ن (ن−١) كمايلى:

$$ر_ك = \frac{مجـ\ د}{\frac{١}{٢} ن (ن - ١)}$$

$$= \frac{١١}{\frac{١}{٢} ٦ (٦ - ١)} = \frac{١١}{٢٥} = +٠,٧٣٣$$

٦- تدل قيمة معامل الإرتباط (+ ٠,٧٣٣) على أن هناك علاقة طردية متوسطة بين متغيرى الارتفاع وكمية المطر.

مثال (١٠):

نعود إلى المثال رقم (٨) ونتخذ من بياناته الخاصة بمتغيرى معدل النمو السكانى والنسبة المئوية لما يخص الفرد من الإنتاج القومى فى ١٤ دولة مختلفة من دول العالم أساساً لحساب معامل ارتباط الرتب لكندال.

١- نرتب الدول ترتيباً عادياً (تصاعدياً) بلانسبة لمتغير معدل النمو السكانى (س$_{١}$).

٢- ننظم ترتيب الدول بالنسبة للمتغير الثانى (نصيب الفرد من الدخل القومى ش$_{٢}$) مقابل ترتيب معدل نموها السكانى كمايلى:

(س$_{١}$)	١	٢	٣	٤	٥	٦	٧	٨	٩,٥	١١	١٢	١٣	١٤
(ش$_{٢}$)	٥,٥	١٠	١٢,٥	١٤	١٠	١٢,٥	٧	٢	٤	٥,٥	١	٣	٨

٣- وكما فعلنا فى المثال السابق نبدأ بملاحظة ترتيب المتغير (س$_{٢}$) الذى لم يترتب ترتيباً عادياً، ونسجل لكل ترتيب اختلافه عن الترتيب الذى يليه بإعطاء القيمة (+١) للرتب الأكبر، والقيمة (-١) للرتب الأصغر، ثم نجمع الناتج مبتدئين بالترتيب ٥,٥ الذى هو أول ترتيب للمتغير (س$_{٢}$). ونظراً لوجود بعض الرتب المتكافئة (المتساوية فى الترتيب) فإننا نعطى الرتب التى تقع على يسار الرتبة المتسارية معها فى الترتيب القيمة (صفر). فمثلاً الرتبة الأولى ٥,٥ تقع على يسارها الرتب الأكبر والأصغر منها وبينها رتبة أخرى مساوية لها هى ٥,٥ وتعطى للأخيرة القيمة (صفر). كذلك نلاحظ أن الرتبتين ٩,٥ ، ٩,٥ للمتغير (س$_{١}$) تقابلان من رتب المتغير الثانى (س$_{٢}$) الرتبة ٣ والرتبة ٤، وفى هذه الحالة فإن موضع رتب المتغير الثانى تعتمد اعتماداً كبيراً عل يموضع الرتب المتكافئة للمتغير الأول، ولكن يصعب تحديد أى من رتبتى المتغير الأول يجب أن تقابل الرتبة ٢ أو تقابل الرتبة ٤ من رتب المتغير الثانى. وهذا التحديد

له أهمية كبيرة فى تحديد المجموع الكلى (د) للاختلاف بين الرتب. فمثلاً إذا وضعنا الرتبة ٤ قبل الرتبة ٢ ، لزاد المجموع الكلى (د) ٢ . وللتغلب على هذه المشكلة تبقى الرتبة الأولى من المتغير (س٢) التى تقابل إحدى الرتب المتكافئة من رتب المتغير (س١) كما هى فى وضعها وبنفس ترتيبها، بينما يعطى للرتبة الثانية من المتغير (س٢) التى تقابل الرتبة المتكافئة الأخرى من رتب المتغير (س) القيمة صفر. فمثلاً إذا كانت الرتبة ٢ هى الرتبة الأولى من رتب (س٢) والتى سنقابل الرتبة المتكافئة الأولى من رتب (س١) فإن الرتبة ٤ تأخذ القيمة (صفر) عند حساب المجموع الكلى لاختلاف الرتب (د). وهكذا إذا كانت الرتبة ٤ هى الأولى فإن الرتبة ٢ من رتب (س٢) تأخذ القيمة (صفر)، وذلك لأنه فى كل من الحالتين يكون (+١) + (−١) = صفر.

٤– يحسب المجموع الكلى للفروق بين رتب س١ ، س٢ السابقة على النحو التالى:

+ (٧−) + (٧−١) + (٩−) + (٨−١) + (٧−٣) + (٧−٣) + (٤−٨) = (د).

٣٣− = (١+) + (٢+) + (٢−١) + (٢−٢) + (١−٣) + (٥−١)

وبالتعويض عن (د) بالقيمة (−٣٣) و ن = ١٤ فى معادلة كندال لحساب قيمة معامل ارتباط الرتب، نجد أن:

$$\text{ر ك} = \frac{\text{د}}{\frac{1}{2}\,\text{ن}\,(\text{ن} - 1)}$$

$$= \frac{-33}{\frac{1}{2}\,(14)\,(14-1)} = -0{,}363$$

وكما هو واضح فإن قيمة معامل ارتباط كندال المحسوبة (−٠,٣٦٣) أقل بكثير من معامل ارتباط الرتب السابق لنفس البيانات.

الطريقة الثانية:

نفرض – على سبيل المثال – أننا نريد حساب معامل كندال لارتباط الرتب للمثال رقم (٢) الخاص بطول النهر (س$_{١}$) ومساحة حوضه (س$_{٢}$) فنتبع الخطوات التالية.

١– نقوم بإعادة ترتيب رتب أحد المتغيرين حسب الترتيب العادى (١، ٢، ٣، ...) فمن المثال السابق الذى يبين طول النهر (س$_{١}$) ومساحة حوضه (س$_{٢}$) نحصل على:

س$_{١}$:	٩,٤٠	١٠,٨٥	١٠,٦٩	١٠,٩٧	١١,٩٤	١٨,٦٩	١٧,٢٣	١٨,٤١	٢٠,١٤	٢٠,٦٨
س$_{٢}$:	٤,٠٨	٤,٥٣	٤,٩٢	٦,١٨	٦,٢٢	٦,٥٤	٧,٥٥	٨,٠٢	١٠,١٦	١٢,٥٠
ترتيب س$_{١}$:	١	٣	٢	٤	٥	٨	٦	٧	٩	١٠
ترتيب س$_{٢}$:	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠

٢– نبدأ بملاحظة المتغير الذى لم يترتب ترتيباً عادياً (المتغير س$_{١}$) بادئين بالرتبة الأولى (وهى ١) ونبحث عن الرتبة العادية المناظرة لها فى المتغير س$_{٢}$ فنجدها ١ وهذا الترتيب على يساره ٩ رتب ولا توجد رتباً على يمينه، وعلى ذلك يكون فرق عدد الرتب = (٩– صفر) = +٩، ثم نشطب الرتبة ١ من الرتب العادية فنحصل على:

٣	٤	٥	٨	٦	٧	٩	١٠
٢	٤	٥	٦	٧	٨	٩	١٠

ونبحث عن الترتيب المناظر للرتبة ٢ فى المتغير (س$_{١}$) فنجدها ٣، ونجد على يسارها ٧ رتب وعلى يمنها توجد رتبة واحدة وبذلك يكون فرق عدد الرتب = (٧–١) +٦، ثم نشطب ٢ من الرتب العادية فنحصل على:

٣	٤	٥	٨	٦	٧	٩	١٠
٢	٤	٥	٦	٧	٨	٩	١٠

فنجد أن ٢ (وهو الترتيب المناظر للترتيب ٣) يقع على يساره ٧ رتب وعلى يمينه لاتوجد أية رتب، فيكون الفرق فى عدد الرتب هو ٧- صفر = +٧. وهكذا، فتكون الفروق فى عدد الرتب هى:

+ ٩، +٦، + ٧، + ٦، + ٥، + ٢، + ١، + ٢، + ١، صفر

ويكون مجموع الفروق فى عدد الرتب = ٣٩

٣- نحسب معامل كندال لارتباط الرتب من المعادلة:

$$ر_ك = \frac{٢ \text{ مجـ } د}{ن (ن - ١)}$$

$$= \frac{٢ \times ٢٩}{١٠ (١٠ - ١)} = \frac{٧٨}{٩٠}$$

$$= + ٠,٨٨٦$$

وكما هو واضح فهو قريب من معامل ارتباط الرتب السابق لسبيرمان.

اختبار المعنوية الاحصائية للارتباط:

سبق أن قلنا أن تحليل الإرتباط ماهو إلا وصف إحصائى لدرجة ترابط وعلاقة المتغيرين قيد التحليل، كما ذكرنا أن قيمة معامل الإرتباط المحسوبة بإحدى الطرق الإحصائية المختلفة لاتدل دلالة أكيدة على وجود علاقة بين المتغيرين، إذ قد تكون النتائج التى نحصل عليها متأثرة بعامل الصدفة الناتج من خطأ فى أسلوب المعانية. وعلى ذلك يجب عمل اختيار لقيمة معامل الإرتباط للتأكد به من درجة احتمال أن الإرتباط لا يحدث بطريق الصدفة، أو بمعنى آخر نتعرف به على مدى معنوية هذا الإرتباط وتأثير حجم العينة أو البيانات موضع البحث والتحليل.

ولاختبار معاملات الإرتباط السابقة يوضع فرض العدم القائل أن قيمة الإرتباط بين المتغيرين هى صفر، أو بعبارة أخرى لا يوجد ارتباط بين المتغيرين (أى أنهما مستقلين عن بعضهما البعض). ولاختبار هذا الفرض فاننا نعين مستوى الدلالة أو

المعنوية Significance Level المطلوب سواء لمستوى احتمال ٠,٠٥ أو ٠,٠١ ثم نقوم بمقارنة قيمة معامل الإرتباط المحسوبة بالقيم النظرية فى الجداول الخاصة بكل نوع من أنواع معاملات الإرتباط الثلاثة (راجع ملاحق الجداول الاحصائية بنهاية الكتاب). فإذا كانت قيمة معامل الإرتباط المحسوبة أكبر من نظيرتها فى الجدول بدرجة حرية مساوية لعدد أزواج القيم المشاهدة مطروحاً منها ٢ فإن هذا يعنى رفض فرض العدم وقبول الفرض البديل وهو أنه يوجد فعلاً ارتباط بين المتغيرين (أى أن الإرتباط له دلالة إحصائية عند مستوى المعنوية المطلوب). وسنوضح فيمايلى كيفية إختبار كل نوع من معاملات الإرتباط على حدة.

أولاً: بالنسبة لمعامل ارتباط بيرسون يستخدم توزيع ستيودنت – ت لاختبار قيمة معامل الإرتباط المحسوبة، وذلك بعد تعديل الصيغة إلى:

$$\text{ت} = \frac{\text{ر}\sqrt{\text{ن} - ٢}}{\sqrt{١ - \text{ر}^{٢}}} \quad \text{أو} \quad \text{ر}\,\frac{\sqrt{\text{ن} - ٢}}{\sqrt{١ - (\text{ر})^{٢}}} \qquad .... (١١-٩)$$

حيث ر هى قيمة معامل ارتباط بيرسون، ن هى عدد أزواج القيم ففى المثال رقم (٢) الذى عرضناه عن طول النهر ومساحة حوضه وجدنا أن ر = ٠,٩٠٢ وأن ن = ١٠ وبالتعويض فى الصيغة السابقة نحصل على

$$\text{ت} = \frac{٠,٩٠٢\sqrt{١٠ - ٢}}{\sqrt{١ - (٠,٩٠٢)^{٢}}}$$

وحيث أن قيمة ت النظرية بدرجات الحرية ١٠–٢ = ٨ هى ٢,٣١ لمستوى دلالة ٠,٠٥ أصغر من قيمة ت المحسوبة، فإن ذلك يعنى أننا نرفض فرض العدم القائل أنه لا يوجد اختلاف بين القيمة المحسوبة لمعامل الإرتباط والصفر (الذى يعنى عدم وجود علاقة بين المتغيرين)، ونقبل الفرض البديل وبمعنى آخر أن هناك اختلاف جوهرى بين معامل الإرتباط المحسوب ومعامل ارتباط صفر، وأن قيمة معامل الإرتباط المحسوبة لها دلالة (معنوية) إحصائية عند مستوى ٠,٠٥.

وفى المثال رقم (٥) الذى يختص بحساب قيمة معامل ارتباط بيرسون بين كمية الأمطار وشدة سريان المياه فى أحد الأنهار فإنه بالتعويض فى الصيغة السابقة نحصل على:

$$ت = ٠,٩٢١ \sqrt{\frac{١٦ - ٢}{١ - (٠,٩٢١)^{٢}}} = ٨,٨٤٦$$

وحيث أو قيمة ت النظرية بدرجات الحرية ١٦-٢ = ١٤ هى ٢,١٥ لمستوى دلالة ٠,٠٥ أصغر من قيمة ت المحسوبة، فإن ذلك يعنى أن معامل الإرتباط المحسوب له دلالة (معنوية) إحصائية عند مستوى ٠,٠٥.

ثانياً: فيما يختص بمعامل سبيرمان لارتباط الرتب فإنه يمكن استخدام بيانات الجدول الاحصائى الخاص به (راجع الجداول الإحصائية بنهاية الكتاب) لاستخلاص القيمة المتوقعة لمعامل الإرتباط حسب حجم العينة (ن) ومستوى الدلالة المطلوب. ففى حالة المثال رقم (٨) الخاص بحساب الإرتباط بين معدل النمو السكانى ونصيب الفرد من الإنتاج القومى، نجد أن حجم العينة (عدد الدول) ١٤، وأن مستوى الدلالة هو ٠,٠٥ (أى احتمال أن ٥٪ من الحالات تكون فيها قيمة معامل الإرتباط راجعة إلى الصدفة)، والقيمة المتوقعة التى نحصل عليها من الجدول الاحصائى هى ٠,٤٥٦ وهى قيمة تقل عن القيمة التى حصلنا عليها فى مثالنا وهى -٠,٥٥٥ أى أننا نستطيع أن تنفى بأن لقيمة التى حصلنا عليها ترجع إلى الصدفة بمستوى الدلالة المطلوب، أو بعبارة أخرى أن هناك احتمال مقدارة ٩٥٪ فإن لاتكون قيمة معامل الإرتباط التى حصلنا عليها قد حدثت بفعل الصدفة أو العشوائية فى التوزيع، وأن الإرتباط (العلاقة) بين المتغيرين له دلالة إحصائية عند مستوى المعنوية المطلوب.

كما يمكن استخدام الصيغة الاحصائية المعدلة لاختبار ستيودنت - ت لمعايره قيمة معامل ارتباط سبيرمان المحسوبة لتحديد دلالة أو معنوية هذه القيمة بنفس الطريقة التى اتبعناها عند معايرة قيمة ارتباط بيرسون.

ثالثاً: أما فى حالة معامل ارتباط كندال فإن التعرف على دلالة أو معنوية قيمة المعامل يختلف حسب عدد المفردات. فإذا كان عدد المفردات ١٠ فإننا نلجأ إلى الجداول الإحصائية الخاصة بهذا المعامل (انظر ملحق الجداول الاحصائية) والتى توضح درجة الاحتمال المرتبطة بعدد مفردات العينة (ن). ففى المثال رقم (٩) الخاص بمعرفة درجة الترابط بين كمية المطر والارتفاع عن سطح البحر كان عدد مفردات العينة ٦ والمجموع الكلى للفروق بين هذين المتغيرين (د) هو ١١ فتصبح درجة الاحتمال ٠,٠٢٨ (٣٪ تقريباً) أى أن هناك ٣٪ من الحالات يكون فيها الإرتباط قد حدث بفعل الصدفة، وبعبارة أخرى أن قيمة معامل الإرتباط التى حصلنا عليها وهى ٠,٧٣٣ ذات دلالة إحصائية عند مستوى معنوية ٠,٠٥.

أما إذا كان عدد المفردات للمتغيرين أكثر من ١٠، كما هى الحال فى المثال رقم (١٠) الخاص بمعرفة درجة الترابط بين معدل النمو السكانى ونصيب الفرد من الإنتاج القومى فى بعض دول العالم، فإننا نستخدم توزيع (ز) لمعايرة قيمة معامل الإرتباط المحسوبة لتحديد دلالة أو معنوية هذه القيمة، ويتم ذلك بالمعادلة الآتية:

$$(ز) = \frac{ر_{ك}}{\sqrt{\frac{٢(٢ن + ٥)}{٩ن(ن - ١)}}} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (١١-١٠)$$

وتحدد درجات الحرية على أساس أنها تساوى (ن -٢).

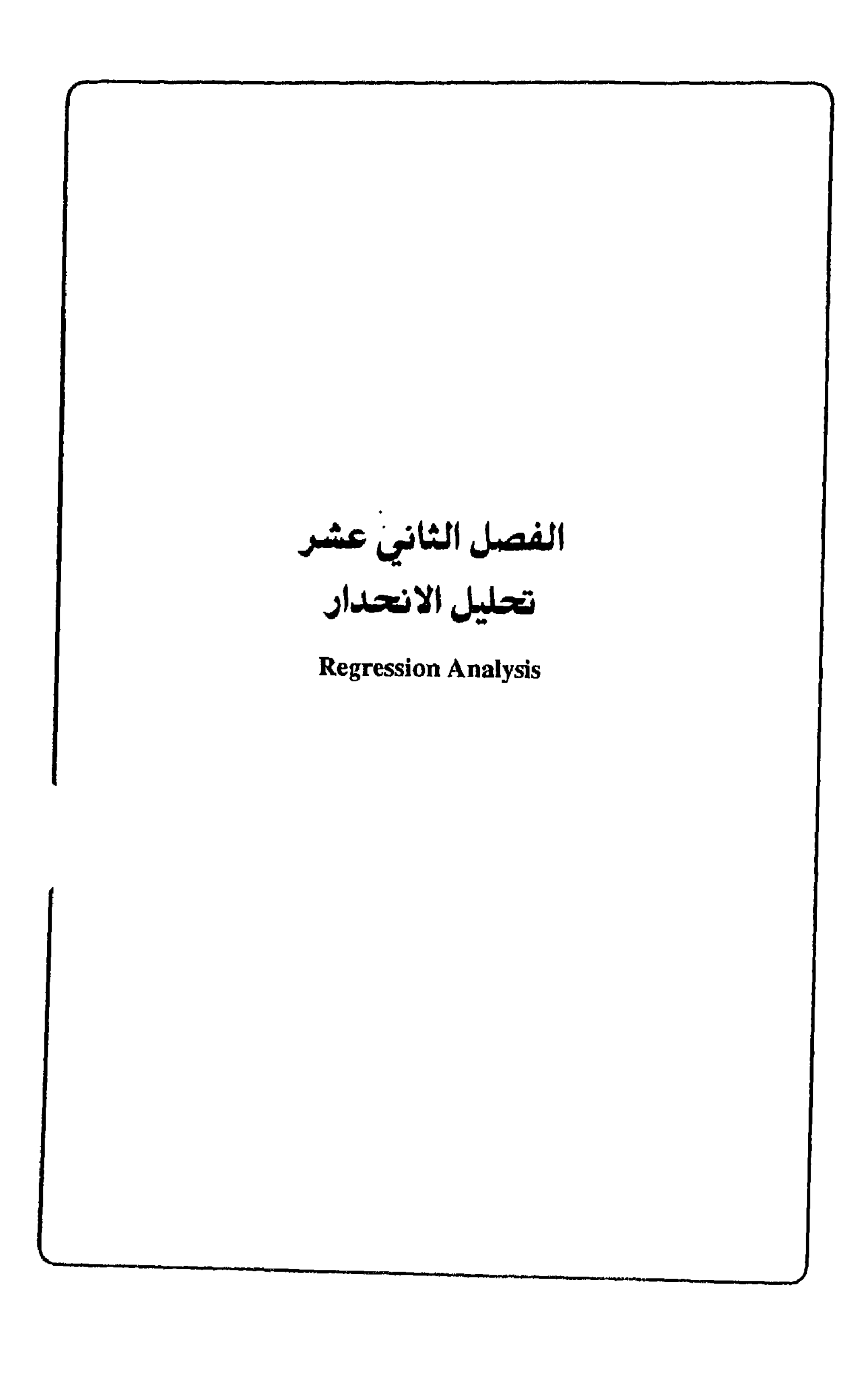
ويمكن حساب قيمة (ز) لمعامل كندال لارتباط الرتب الذى حصلنا عليه من المثال رقم (١٠) كمايلى:

$$(ز) = \frac{-٠,٣٦٣}{\sqrt{\frac{٢(٣٣)}{١٢٦ \times ١٣}}} = \frac{-٠,٣٦٢}{\sqrt{\frac{٦٦}{١٦٣٨}}}$$

$$= -١,٨١$$

وبالرجوع إلى جداول توريع (ر) لنطريه نجد أن قيمة (ز) المحسوبة وهى ١,٨١ لها احتمال ٠٣٦ · ٠ أى أن قيمة معامل كندال المحسوبة لاترجع إلى الصدفة فى مستوى معنوية ٠٥ · ٠ وهذا يعنى بشكل آخر أن هناك احتمال مقداره ٩٥٪ بأن لا يكون الإرتباط فى المثال السابق قد حدث بفعل الصدفة، لأن عامل الصدفة يتدخل فى ٦ ٣٪ من الحالات. وعلى العموم فإنه مع معامل ارتباط كندال نجد أن الاحتمالات الخاصة بقيم د (الفرق بين الرتب) المحسوبة لعدد المفردات (ن) التى تتراوح بين ١٠ ,٤٠ تقل كلما زادت قيمة د، وبالتالى يتخذ ذلك دليلاً على صغر فرصة حدوث الإرتباط بالصدفة.





الفصل الثاني عشر

تحليل الانحدار

Regression Analysis



تحليـــل الإنحـــــدار
Regression Analysis

ذكرنا فى الفصل السابق عن تحليل الإرتباط أن الهدف من حساب معامل الإرتباط هو معرفه درجه العلاقه أو مقدار الترابط بين متغيرين (ظاهرتين). إلا أنه إذا ما وجدت علاقة بين متغيرين فإننا ربما نحتاج إلى التوقع (أو التنبؤ prediction) بسلوك أحد المتغيرين فى ضوء تأثره بمتغير آخر أو بعدة متغيرات أخرى ، أو إذا كانت هناك رغبة فى تقدير مدى تأثير كل متغير من المتغيرات على متغير آخر. وواضح أن مثل التقدير يزداد دقة كلما كان الإرتباط شديداً. ويسمى المتغير الذى يراد دراسة سلوكه ومعرفة مدى تأثره بالمتغيرات الأخرى المتغير التابع (ص) Dependent Veriable ويطلق على التتغير الذى يؤثر فى سلوك المتغير التابع بالمتغير المستقل Independent Veriable. وفى بعض الأحيان يسمى المتغير التابع باسم المتغير المتنبأ له Predictand Veriable أو المتغير المعيار Criterion Veriable، كما أنه يطلق على المتغير المستقل اسم المتغير المتنبيء به Predictor Veriable أو المتغير المفسر Explanatory Veriable.

كما وقد سبق أن أوضحنا أنه لا يشترط لدراسة الإرتباط بين متغيرين أن تكون هناك علاقة دالية بينهما ، ولكن إذا كانت العلاقة الدالية بين متغرين يمكن وصفها إحصائيا بخط مستقيم سميت هذه العلاقة «بعلاقة انحدار خطية» lincar Regression. وبذلك يختص الإنحدار (أو ما يعرف أحيانا بالارتداد أو الاعتماد)

بدراسة العلاقة بين متغيرين على هيئة علاقة دالية بحيث يمكن التنبؤ منها عن أحد المتغيرين بمعلومية المتغير الآخر.

أهداف تحليل الإنحدار:

يستخدم تحليل الانحدار كأسلوب احصائى كمى فى النواحى التالية :

١- تقدير العلاقة بين متغيرين على شكل علاقة دالية: ص= (س) أو س= (ص) والتى عن طريقها يمكن معرفة التغير فى أحد المتغير على أساس تأثره بالمتغير الآخر. أو بعبارة آخرى توقع وتنبؤ سلوك المتغير التابع فى ضوء تأثره بالمتغير أو المتغيرات المستقلة.

٢- قياس مدى الإرتباط الكلى بين المتغير التابع والمتغير أو المتغيرات المستقلة.

٣- تقدير نسبه تفسير كل متغير مستقل للاختلاف فى المتغير التابع.

٤- إجراء سلسلة من الاختبارات الفرضية لأى من العلاقات الثلاثة السابقة.

أنواع تحليل الإنحدار:

هناك ثلاثة أنواع رئيسية لتحليل الإنحدار نجملها فيما يلى:

١- الإنحدار البسيط Simple Regression و هو يستخدم لدراسة العلاقة بين متغيرين فقط.

٢- الإنحدار الجزئى Partial Regrenion و هو يدرس العلاقة بين المتغير التابع وواحد فقط من المتغيرات المستقلة بفرض أن للعوامل الأخرى ثابتة (أى بإهمال تأثير العوامل الأخرى).

٣- الإنحدار المتعدد Multiple Regression وهو يحدد مقدار العلاقة بين المتغير التابع والنتغيرات المستقلة كلها. وهناك نوع آخر مشابه للإنحدار المتعدد يسمى بالإنحدار التدرجى Stepwise Regression وهو يعطى نسبة تفسير كل متغير مستقل فى اختلاف المتغير التابع، وتكون نسبة التفسير مرتبة حسب أهمية

كل متغير مستقل (أى أنه يبدأ بتحديد أعلى نسبة أو أهم متغير وينتهى بأقل نسبة أو المتغير أقل أهميه فى تفسير الاختلاف الذى يحدث فى المتغير التابع).

ورغم وجود بعض الاختلافات فى العمليات الحسابية لكل نوع من أنواع التحليل السابقة إلا أنها تسير فى نفس الاتجاه تقريباً، و هو تحديد معادلة خط الانحدار للعلاقة بين متغيرين أو عدة متغيرات. وسنقتصر فى دراستنا فى هذا الفصل على تحليل النوع الأول من الإنحدار (تحليل الإنحدار البسيط) كأحد التحليلات الشائعه الاستخدام، حيث أن النوعين الآخرين (الانحدار الجزئى والانحدار المتعدد) يحتاجان إلى جهد كبير ووقت طويل فى عملياتهما الحسابية كلما زاد عدد الحالات Cases والمتغيرات Variables. و قد ظهرت أهميه الحاسب الآلى Computer كعامل مساعد هام للقيام بمثل هذه العمليات الحسابية الطويلة. وكان لتقدم الدراسات الخاصة بنظم معالجة البيانات أن تعددت البرامج الجاهزة التى تستخدم فى مثل هذا النوع من التحليل الاحصائى. ومن أمثله هذه البرامج برنامج معروف فى مراكز الحاسبات الآلية الكبيرة فى جامعات المملكة المتحدة ضمن مجموعة من البرامج يطلق عليها (S.P.S.S). ويقوم هذا البرنامج بحساب العلاقات بين عدد من المتغيرات المستقلة و المتغير التابع.

تحليل الانحدار البسيط:

تجدر الإشارة قبل الخوض فى تحليل الإنحدار البسيط (بين متغيرين). أن نذكر أن هناك شرطاً أساسياً فى بيانات المتغيرات المستخدمة للتحليل وهو أن تكون هذه البيانات من نوع بيانات الفترة Intervally-Scaled data فى حالة المتغير التابع، ويستحسن أن تكون كذلك للمتغير المستقل. ولكن أحيانا يمكن أن تكون بيانات المتغير المستقل من نوع البيانات النوعيه Nominally-Scaied data. كما لابد أن نوضح أن هناك بون شاسع بين تحليل الإرتباط و تحليل الإنحدار. فعلى الرغم من تشابه العلاقات الرياضية بين الإرتباط و الإنحدار إلا أنهما يختلفان عن بعضهما فى النواحى التالية:

١- تحليل الإرتباط عبارة عن مقياس وصفي بينما تحليل الإنحدار مقياس كمي.

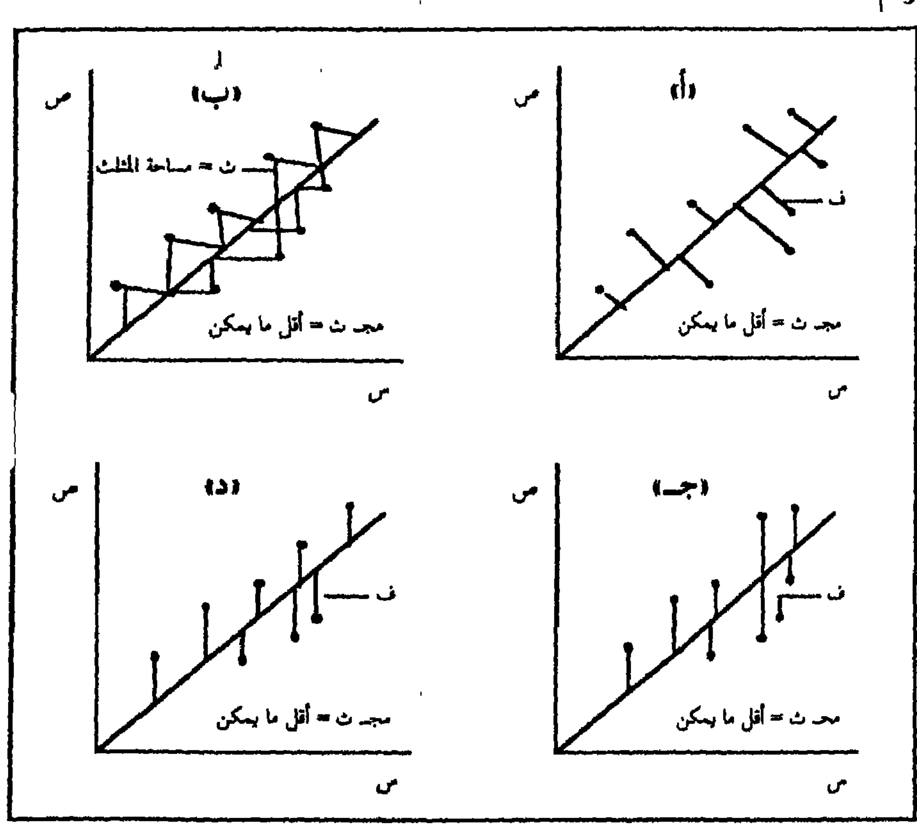
٢- يشترط في تحليل الإرتباط أن يكون توزيع بيانات كل المتغيرات توزيعاً معتدلاً، بينما يشترط في تحليل الإنحدار أن تكون بيانات المتغير التابع (ص) فقط ذات توزيعاً معتدلاً ، أما المتغير المستقل (س) فيجب أن تكون قيم مفرداته ثابتة Fixid (أى أن قيم (س) تقاس بدون أخطاء)، ولو أن الإنحدار يتأثر بوحدات القياس - إلا أنه في بعض الحالات التي يحسب فيها الإرتباط والإنحدار فإننا نتغاضى عن الشرط الخاص بأن قيم المتغير المستقل (س) تكون قيماً ثابتة.

٣- يشترط في تحليل الإنحدار أن تكون هناك علاقه دالية بين المتغيرات ، بينما لايشترط ذلك في تحليل الإرتباط.

و بصفة عامة فإن استخدام أسلوب الإرتباط لايحقق سوى قياس درجة العلاقة الإرتباطية بين المتغيرين، بينما الهدف من أسلوب الإنحدار هو دراسة التوقع أو التنبؤ بتغير المتغير التابع في ضوء معرفة التغيرات في المتغير المستقل وهو ما يطلق عليه بالعلاقة الدالية.

و يعتمد تحليل الإنحدار لدراسة العلاقة بين ظاهرتين على تكوين فكرة مبدئية عن نوع هذه العلاقة وقوتها وذلك باستخدام ما يعرف بشكل الانتشار Scatter Diagram. فإذا مثلنا أزواج المشاهدات (القيم) الخاصة بالظاهرتين بيانياً نحصل على عدد من النقط التي قد تقع تماماً على خط مستقيم فيكون الإرتباط تاماً - أو قد تنحرف عنه أو تأخذ شكلاً آخر غير الخط المستقيم، وعلى العموم إذا كانت هناك علاقة تربط الظاهرتين فإن النقط تنتشر بشكل منتظم حسب نوع العلاقة الموجودة (علاقة عكسية أو علاقة طردية). أما إذا كانت النقط مبعثرة دون نظام ملحوظ فإن العلاقة تكون ضعيفة جداً أو منعدمة (أنظر الفصل السابق عن تحليل الإرتباط). والخط الذى تنتشر حوله النقط بانتظام يسمى خط الانتشار أو خط الإنحدار (ويكون هذا الخط مستقيماً أو منحنياً).

ومن المعلوم أن الخط الذى نوفقه لا يمر بجميع النقط (إلا فى حالات خاصة) فى شكل الإنتشار وعلى ذلك تكون هناك بعض النقط التى تنحرف عن هذا الخط، وبالتالى إذا اخترنا أى قيمة للمتغير المستقل بمعلومية احداثيها الأفقى و قدرنا قيمه الاحداثى الرأسى للمتغير التابع المناظرة لها فإن القيمة الأخيرة (المقدر) سوف تختلف عن قيمه الاحداثى الرأسى المشاهدة و الفرق بين القيمتين المقدره والمشاهدة يسمى إنحراف النقطة (البعد الرأسى لها) عن الإنحدار (شكل رقم ١٢- ١)



شكل رقم (١٢-١)

انحراف نقط تمثيل المتغيرين س ، ص عن خط الانحدار

و يمكن حساب معادلة خط الإنحدار بطريقة المربعات الصغرى التى من

خصائصها أن يكون مجموع مربعات انحرافات النقط عن خط الإنحدار أصغر ما يمكن. ويكون خط الإنحدار مستقيماً إذا كانت العلاقة بين المتغيرين علاقة دالية. والمعروف أن معادلة الخط المستقيم هي معادلة من الدرجة الأولى على صورة:

$$\text{ص} = \text{م س} \pm \text{ج}$$

حيث ص في هذه الحاله هي قيمه المتغير التابع، س هي قيمه المتغير المستقل، وحيث م مقدار ثابت يمثل ميل خط الانحدار على المحور الافقي، ويسمى معامل الإنحدار Regression Coefficeent، و جـ مقدار ثابت أيضاً هو طول الجزء المقطوع من المحور الرأسي بواسطة خط الإنحدار. و بتعين م ، جـ يتعين خط الإنحدار كما يمكن تقدير قيمة (ص) و هي القيمة التي تمثل التوقع أو التنبؤ المطلوب حيث أن قيمة (س) معروفة.

ولإيجاد معادلة خط الإنحدار على الصورة السابقة نحسب قيم م ، جـ التي تحقق الشرط العام لهذا الخط وهو أن مجموع مربعات انحرافات الأبعاد الرأسية للنقط عنه تكون أصغر ما يمكن. ويمكن الحصول عليها بواسطة عدة طرق أكثرها شيوعاً الطرق التالية:

$$(١)\quad \text{م} = \frac{\text{مجـ س ص} - \dfrac{(\text{مجـ س} \times \text{مجـ ص})}{\text{ن}}}{\text{مجـ س}^{٢} - \dfrac{(\text{مجـ س})^{٢}}{\text{ن}}}$$

$$(٢)\quad \text{م} = \frac{\dfrac{\text{مجـ س ص}}{\text{ن}} - \bar{\text{س}}\,\bar{\text{ص}}}{\text{ع}^{٢}\,\text{س}}$$

حيث سَ المتوسط الحسابي للمتغير المستقل، صَ هي المتوسط الحسابي للمتغير التابع، ع^{٢} س هي تباين المتغير المستقل.

$$(٣)\quad \text{م} = \frac{\text{مجـ س ص} - \text{ن } \bar{\text{س}}\, \bar{\text{ص}}}{\text{مجـ س}^{٢} - \text{ن } \bar{\text{س}}^{٢}}$$

حيث أن هى عدد أزواج القيم للمتغيرين.

$$(٤)\quad \text{م} = \frac{\text{مجـ } (\text{س} - \bar{\text{س}})\,(\text{ص} - \bar{\text{ص}})}{\text{مجـ } (\text{س} - \bar{\text{س}})^{٢}}$$

إلا أنه لتسهيل العمليات الحسابية فإنه يفضل استخدام الصيغة رقم (٣). أما قيمة (جـ) فتحسب بالطريقة التالية بعد معرفة قيمة (م).

$$\text{جـ} = \bar{\text{ص}} - \text{م } \bar{\text{س}}$$

وعن طريق معرفة قيمة كل من م، جـ يمكن تحديد قيم المتغير التابع (ص) فى ضوء أى تغير فى قيم المتغير المستقل (س). ويرمز لقيم (ص) المتوقعة بالرمز $\hat{\text{ص}}$. وتكون المعادلة المستخدمة فى التوقع أو التنبؤ الصحيح فى حالة متغيرين فقط كالآتى:

$$\text{ص}_{١} = \text{م س} + \text{جـ}$$

مثال:

فى المثال الخاص لدراسة العلاقة بين مساحة حوض النهر وطوله (مثال رقم: ٢ – الفصل الرابع عشر) فإنه لإيجاد المعادلة الخطية العلاقة بين هذين المتغيرين نقوم بوضع البيانات فى جدول ونحسب منه قيمة (ص) وذلك على النحو التالى:

جدول رقم (١٢ - ١)

مساحة حوض النهر وطوله

المتغير المستقل س (طول النهر الكلى)	المتغير التابع ص (مساحة حوض النهر)	$س^2$	س ص
٩,٤٠	٤,٠٨	٨٨,٤	٣٨,٤
١٠,٨٥	٤,٥٣	١١٧,٧	٤٩,٢
١٠,٦٩	٤,٩٢	١١٤,٣	٥٢,٦
١٠,٩٧	٦,١٨	١٢٠,٣	٦٧,٦
١١,٩٤	٦,٢٢	١٤٢,٦	٧٤,٣
١٨,٦٩	٦,٥٤	٣٤٩,٣	١٢٢,٢
١٧,٢٣	٧,٥٥	٢٩٦,٩	١٣٠,١
١٨,٤١	٨,٩٢	٣٣٩,٩	١٦٤,٢
٢٠,١٤	١٠,١٦	٤٠٥,٦	٢٠٤,٦
٢٠,٦٨	١٢,٥٠	٤٢٧,٧	٢٥٨,٥
١٤٩,٠٠	٧١,٦	٢٤٠٢,٧	١١٦١,٩

سَ = ١٤,٩   صَ = ٧,١٦

وبتطبيق الصيغة رقم (٣) لحساب قيمة (م) فإن:

$$م = \frac{١١٦١,٩ - ١٠ \times ٧,١٦ \times ١٤,٩}{٢٤٠٢,٧ - ١٠ \,(١٤,٩)^{٢}}$$

$$= \frac{٩٥,٠٦}{١٨٢,٦} = + ٠,٥٢$$

(لاحظ أن م موجبة لأن العلاقة طردية) ولإيجاد قيمة جـ (صَ - م سَ) فإن:

جـ = ٧,١٦ - ,٥٢ × ١٤,٩ = -,٥٨٨

ويمكن حساب قيمة (م) على أساس الصيغة رقم (١)

$$= \frac{١١٦١٩ - \frac{١٤٩,٩٨ \times ٧١.٦}{١٠}}{٢٤٠٢,٧ - \frac{(١٤٩,٩٨)^{٢}}{١٠}}$$

$$= \frac{٩٥,٠٦}{١٨٢,٦} = ٠,٥٢$$ (وهى نفس النتيجة التى حصلنا عليها سابقاً)

ويوضح الشكل رقم (١٢ – ٢) كيفية رسم خط الانحدار الذى يمثل العلاقة المتوقعة بين الظاهرتين مساحة حوض النهر وطوله. فبعد أن يتم رسم المحور الأفقى والمحور الرأسى توقع النقط التى تمثل قيم الظاهرتين (المتغيرين س ، ص) على الرسم. فمثلاً النقطة (١) فى الشكل تبين نقطة التقاء قيمة ش المشاهدة (٩,٤٠) وقيمة (ص) المشاهدة (٤,٠٨). وتمثل النقطة (٥) فى الرسم نقطة التقاء س المشاهدة (١١,٩٤) و (ص) المشاهدة (٦,٢٢)، كذلك تمثل النقطة (٧) التقاء قيمة (س) المشاهدة (١٧,٢١) بقيمة (ص) المشاهدة (٧,٥٥) وهكذا. وبعد أن يتم تمثيل جميع النقط يرسم خط الانحدار وذلك بتوصيل نقطتين: الأولى تمثل قيمة (جـ) على محور الصادات وهى فى حالة المثال السابق (–٠,٥٨٨)، والنقطة الثانية تمثل نقطة التقاء قيمتى متوسط س، ص (سَ ، صَ) وهى فى هذه الحالة (١٤,٩٠)، (٧,١٦) وتمثلها النقطة (أ) فى الشكل. وبعد أن توصل هاتين النقطتين بخط يكتب على امتداده المعادلة التى يستدل بها عليه وهى عبارة عن:

ص = ٠,٥٢ س – ٠,٥٨٨

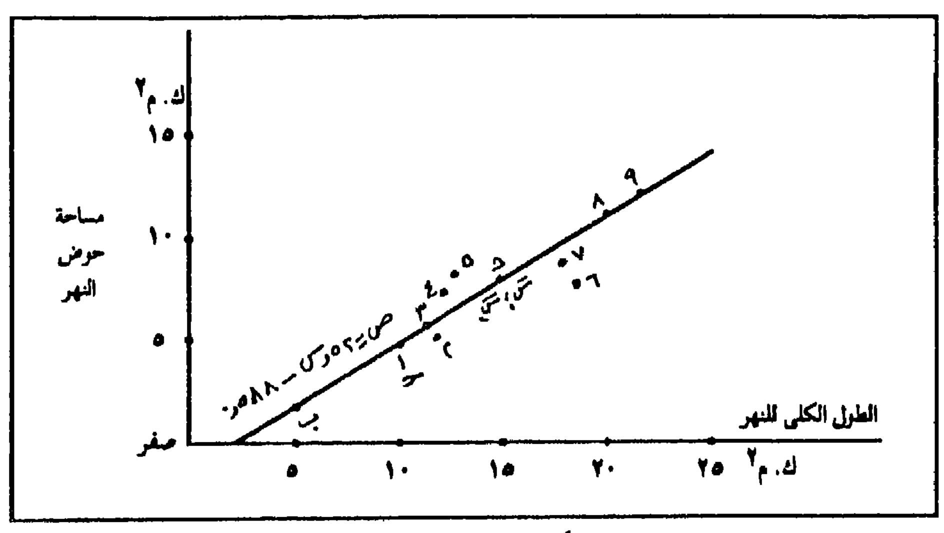
وهناك طريقة أخرى لرسم خط الانحدار بين النقط الموزعة على شكل الانتشار. وهى أن نطبق معادلة خط الانحدار التى حصلنا عليها وذلك لأية قيمتين من قيم (س)، وعن طريقها تحدد قيمتى (ص) المناظرتين وبعد توقيعهما على

الرسم توصل بينهمـا بخط ونمده حتى يلتقى بالمحور الرأسى. ونوضح ذلك عن طريق حساب قيمة (ص) عندما تكون قيمة ٠س) ١٥,٥ كما يلى:

(١) ص = (٠,٥٢ × ٥) − ٠,٥٨٨ = ٢,٠١

(٢) ص = (٠,٥٢ × ٥) − ٠,٥٨٨ = ٧,١٢

وتمثل النقطتان ب، د فى الشكل قيمتى (ص) ٢,٠١، ٧,١٢ المحسوبة لقيمتى (س) ١٥,٥ على الترتيب (لاحظ أن كل من القيمتين تقع على نفس خط الانحدار المرسوم بالطريقة الأخرى).



شكل رقم (١٢ – ٢)

شكل الانتشار وخط الانحدار للعلاقة بين مساحة حوض النهر وطوله

ومن الأهمية فى تحليل الانحدار أن نوضح ما إذا كان معامل الانحدار (م) له دلالة إحصائية تمكن الباحث من معرفة مدى ترابط العلاقة بين المتغيرين (س، ص)، أو بمعنى آخر توضيح ما إذا كان الارتباط بين المتغيرين قد تم بتأثير الصدفة. واختبار مستوى دلالة معامل الانحدار يعتمد على حساب قيمة (ت) حيث يكون الفرض المختبر هو أن معامل الانحدار (م) يساوى صفراً.

ومن المعروف فى تحليل الانحدار أنه يوجد نوعان من الانحرافات: النوع الأول يسمى **الانحرافات عن الانحدار**. وهى عبارة عن الفرق بين متوسط قيم المتغير التابع أو (ص) والقيم المتوقعة (صً) لهذا المتغير والمحسوبة من معادلة الانحدار، أما النوع الآخر من الانحرافات فيطلق عليه الانحرافات العشوائية والتى ترجع إلى الأخطاء العشوائية. ويمكن إجراء اختبار مستوى دلالة معامل الانحدار (م) على النحو التالى:

$$\text{ت} = \frac{\text{م}}{\text{ق}_{\text{م}}}$$ حيث ق م هى الخطأ المعيارى لمعامل الانحدار (م)

$$\text{وقيمة ق م} = \frac{\sqrt{\dfrac{\text{مربع مجموع انحرافات الانحدار}}{\text{عدد المفردات} - ٢}}}{\text{مجموع مربع انحرافات المتغير المستقل عن متوسطه الحسابى}}$$

$$= \frac{\sqrt{\dfrac{(\text{مجـ}[\text{ص} - \hat{\text{ص}}])^{٢}}{\text{ن} - ٢}}}{\text{مجـ}(\text{س} - \bar{\text{س}})^{٢}}$$

وتسمى القيمة $\frac{(\text{مجـ}[\text{ص} - \hat{\text{ص}}])^{٢}}{\text{ن} - ٢}$ تباين الانحدار الذى يحسب من قيم (ص) المتوقعة التى نحصل عليها بتطبيق معادلة خط الانحدار السابقة (راجع صفحة   )، ثم نحسب انحرافات الانحدار كما يلى:

ص (المشاهدة)	صَ (المتوقعة)	(ص − صَ)
٩,٤٠	٤,٣٠	٥,١٠
١٠,٨٥	٥,٠٥	٥,٨٠
١٠,٦٩	٤,٩٧	٥,٧٢
١٠,٩٧	٥,١٢	٥,٨٥
١١,٩٤	٥,٦٢	٦,٣٢
١٨,٦٩	٩,١٣	٩,٥٦
١٧,٢٣	٨,٣٧	٨,٨٦
١٨,٤١	٨,٩٨	٩,٤٣
٢٠,١٤	٩,٨٨	١٠,٢٦
٢٠,٦٨	١٠,١٦	١٠,٥٢
		٧٧,٤٢

$$ق_م = \frac{\sqrt{\frac{(٧٧,٤٢)^{٢}}{٨}}}{١٧٩,٦٤}$$

حيث أن ١٧٩,٦٤ هى قيمة مجـ (س − سَ)٢ من المثال السابق (راجع فصل تحليل الارتباط).

$$= \frac{٢٧,٣٧}{١٧٩,٦٤} = ,١٥٢$$

$$\text{إذن } ت = \frac{م}{ق_م} = \frac{,٥٢}{,١٥٢} = ٣,٤٢$$

وحيث أن قيمة (ت) النظرية فى جداول توزيع ت بمستوى دلالة ٠١ .

بدرجات حرية (ن - ٢) = (١٠ - ٢) = ٨ هى ٣,٣٦، وحيث أن قيمة ت المحسوبة ٣,٤٢ أكبر من قيمة (ت) النظرية لذلك نرفض الفرض القائل بأنه لاتوجد أية علاقة بين المتغيرين (س، ص). أو بعبارة أخرى نقبل الفرض البديل وهو أن هناك علاقة بين المتغيرين (س ، ص) فى مستوى دلالة أو معنوى ٠١, أى أن هناك احتمال قدره ٩٩٪ أن لاتكون العلاقة بين المتغيرين (س ، ص) قد حدثت بفعل الصدفة أو العشوائية.

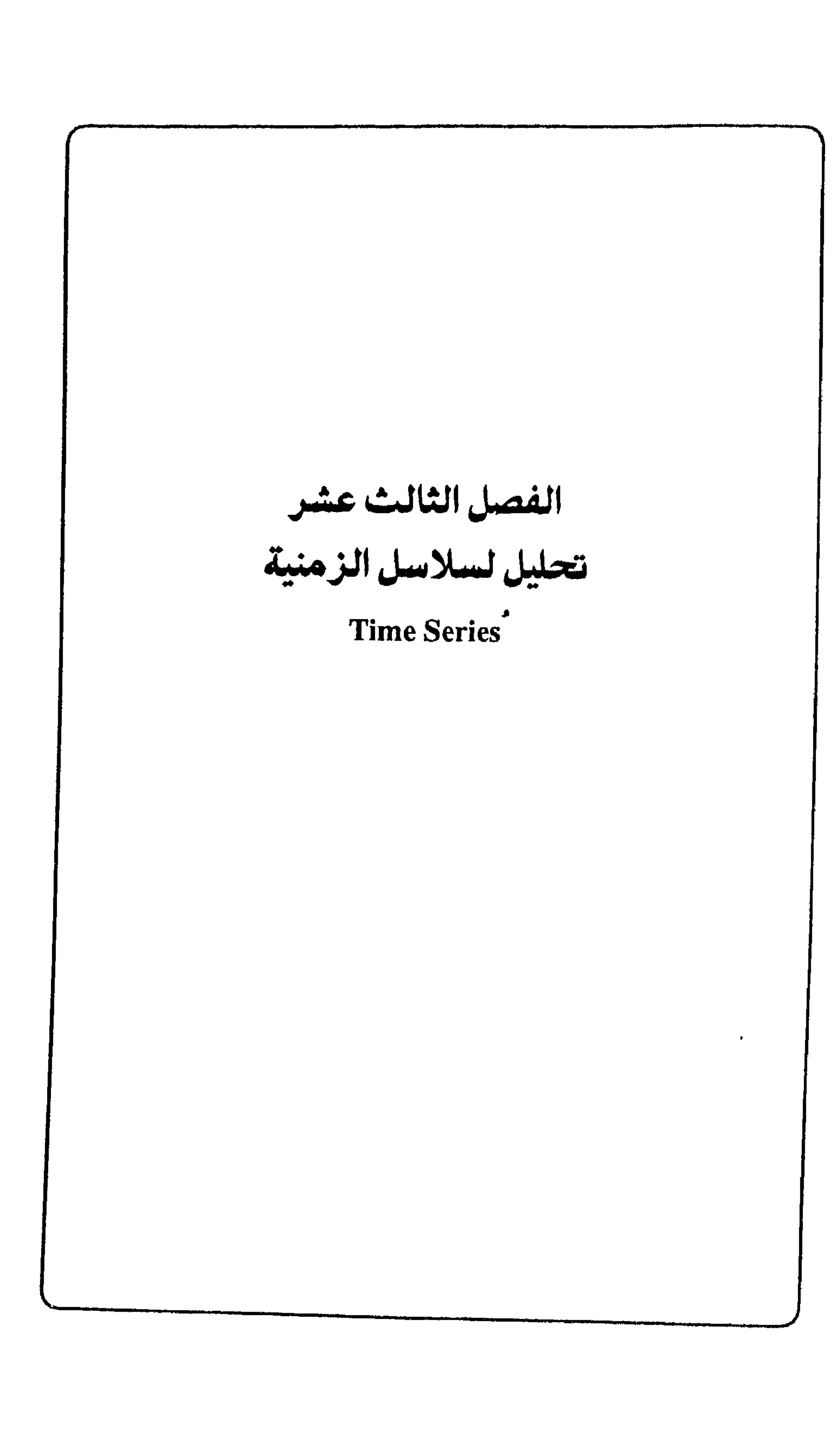
وبعد أن أوضحنا كيفية حساب معادلة خط الانحدار بطريقة المربعات الصغرى، وتمثيلها بيانياً وذلك لتحديد نوع العلاقة بين المتغيرين (س ، ط)، يبقى أن نحدد قوة هذه العلاقة. وتقاس قوة لعلاقة بين المتغير التابع والمتغير المستقل بواسطة ما يسمى بمقياس معامل التحديد Determination coefficient، ويرمز له بالرمز ($ر^2$)، وهو عبارة عن مربع معامل الارتباط فإذا كان معامل الارتباط بين (س،ص) هو ٧٥, مثلاً فإن معامل التحديد يساوى $(٧٥,)^2$ = ٠,٥٦ والفكرة وراء حساب معامل التحديد هى قياس مدى الاختلاف فى قيم ص التى ترجع إلى الاختلاف فى قيم (س)، وبناء عليه إذا كانت العلاقة بين س ، ص قوية فإن ذلك يعنى ارتفاع قيمة معامل التحديد. ونحسب قيمة معامل التحديد $ر^2$ بالمعادلة الآتية:

$$ر^2 = \frac{\text{م [مجـ (س - سَ) (ص - صَ)]}}{\text{مجـ }(\text{ص - صَ})^2}$$

وإذا رجعنا إلى المثال السابق الخاص بالعلاقة بين مساحة حوض النهر وطوله وحسبنا معامل التحديد لهذين المتغيرين لكان هو:

$$ر^2 = \frac{(٩٩,١٣,٥٢)}{٦٧,٢٣} = \frac{٥١,٥٥}{٦٧,٢٣} = ٠,٧٦٧$$

أى أن ٠,٧٧ (٧٧٪) من الاختـلاف فى قـيم (ص) يمكن تفسـيرها بالاختلافات فى قيم «س»، وأن ٠,٣٣ (٣٣٪) من هذه الاختلافات ترجع إلى عامل الصدفة والأخطاء العشوائية.



الفصل الثالث عشر

تحليل لسلاسل الزمنية

Time Series



الفصل الثالث عشر

تحليل السلاسل الزمنية

Time Series

هناك كثير من الظواهر التى تتغير نحو الزيادة أو النقص بمرور الزمن، وإذا ما تتبعنا مشاهدتها فإننا نحصل على سلسلة Series من هذه المشاهدات. غير أننا غالباً ما نجد أن هذه المشاهدات لا تتزايد أو تتناقص باستمرار، ولكنها تتذبذب بين الزيادة إلى النقص أو العكس على حسب الفترات الزمنية المحددة للدراسة. وتسمى مثل هذه السلسلة من المشاهدات بالسلسلة الزمنية. ومن ثم فإن السلسلة الزمنية هى مجموعة من المشاهدات لظاهرة ما أخذت فى فترات زمنية محددة، عادة على فترات أو أبعاد زمنية متساوية (منتظمة). وهى بذلك تحتوى على متغيرين، أحدهما الزمن (المتغير المستقل: ز)، والآخر هو قيمة الظاهرة (المتغير التابع: ص)، وبعبارة أخرى تعرف السلسلة الزمنية رياضياً بالقيم ص١، ص٢، ص٣ ... ص ن والتى يأخذها المتغير ص عند الزمن ز١، ز٢، ز٣ ... ز ن، أى أن ص دالة ز، ويرمز لذلك بالرمز ص = د ز.

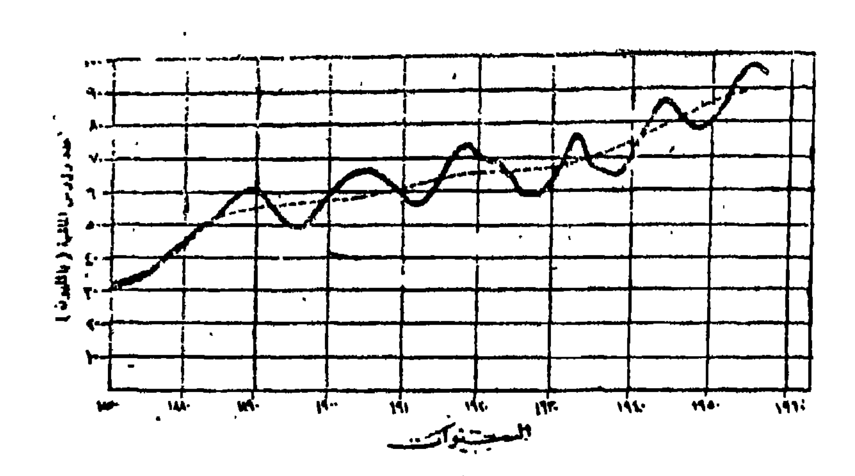
والغرض من دراسة السلسلة الزمنية هو التعرف على التغيرات الكمية التى تطرأ على الظاهرة عبر الزمن لمعرفة أسبابها ونتائجها ومايمكن أن يكون هناك من علاقة بينها وبين غيرها من الظواهر، وكذلك التنبؤ الاحصائى بقيمها غير المشاهدة أو المرصودة، ومالذلك من أهمية فى اتخاذ القرارات التى تتعلق بتخطيط المستقبل. ومن أمثلة السلاسل الزمنية الاستهلاك المحلى فى السنة من محصول القطن فى مصر على مدار عدد من السنوات، الإنتاج الكلى فى السنة من خامات الصلب فى

المملكة المتحدة على فترة زمنية تتكون من عدد من السنوات، كمية الأمطار الساقطة في السنة على أحد المراصد الجوية في مدينة ما لعدة سنوات، عدد السفن التي تعبر قناة السويس على مدار عدد من السنين، عدد سكان مصر في التعدادات المتتابعة .... إلخ.

وسنتعرض في هذا الفصل لدراسة تطور الظواهر الجغرافية على مدار فترات زمنية متتابعة بغرض وصف هذا التطور، والتعرف على طرق قياس التغيرات المختلفة التي تطرأ على هذا التطور بفعل العوامل المتنوعة والمؤثرات العديدة مما يفيد في عملية التنبؤ بما يمكن أن تكون عليه بيانات الظواهر المختلفة في فترات زمنية مضت أولم تأت بعد.

التمثيل البياني للسلاسل الزمنية:

ولغرض قياس التسلسل الزمني يمكن تمثيل السلسلة الزمنية المتضمنة الظاهرة (المتغير ص) مقابل عنصر الزمن (ز) بيانياً كما فعلنا ذلك من قبل (راجع الفصل الثالث). وتعتبر الرسوم البيانية الخطية Line Graph أهم أنواع الرسوم البيانية التي تستعمل لبيان العلاقة بين الظاهرة (المتغير) موضع التحليل ومدى ارتباطها بعنصر الزمن، بحيث توضح أية نقطة على الرسم البياني مقدار هذه الظاهرة خلال فترة زمنية معينة. وهناك نوعان من الرسوم البيانية الخطية يختص النوع الأول منها بالتغيرات الموجبة (نمو مضطرد Growth) والتغيرات السالبة (تناقص أو اضمحلال Decline) لظاهرة معينة، بينما يوضح النوع الثاني التغيرات النسبية Proportional Changes للظاهرة. والنوع الأول يعتمد في تمثيله على التغيرات الحسابية (العددية) Numerical Changes المبنية على الأرقام الفعلية للظاهرة قيد التحليل. وفي هذا النوع من الرسوم يسمى الخط البياني الذى يصل النقط التي تمثل قيم الظاهرة «بالمنحنى التاريخي» Historigram الذى يستخدم لدراسة الاتجاه العام للظاهرة قيد الدراسة (شكل رقم ١٣-١).



شكل رقم (١٣-١)

المنحنى التاريخى لسلسلة أعداد والماشية

فى الولايات المتحدة الأمريكية (١٨٧٠ - ١٩٦٠)ويعاب على هذا النوع من الرسوم البيانية أن دراسة وتحليل الخط البيانى أو المنحنى التاريخى تقودنا إلى نتائج مضللة أحياناً، خاصة إذا كان الغرض من التحليل هو مقارنة هذه النتائج العددية مع غيرها من القيم العددية للظواهر الأخرى.

ولبيان وتمثيل التغيرات النسبية فى السلسلة الزمنية تستخدم طريقة الأرقام القياسية Index Numbers التى تعتمد على اختيار سنة الأساس Base Year ويعطى لها الرقم ١,٠٠ ثم تحول بقية القيم فى السلسلة إلى نسب مئوية منسوبة إلى سنة الأساس المختارة. فمثلاً إذا أردنا تحويل القيم العددية فى الدجلو التالى (جدول رقم ١٣-١) إلى أرقام قياسية نتبع الخطوات التالية:

جدول رقم (١٣– ١)

تطور أعداد رؤوس الماشية والأغنام فى اسكتلندا

(بالألف رأس)

السنة	عدد الماشية	عدد الأغنام	السنة	عدد الماشية	عدد الأغنام
١٩٤٦	١٤٧٢	٦٢٥٤	١٩٥٨	١٨٢٠	٧٩٢٩
١٩٤٨	١٤٩٩	٦٧٣١	١٩٦٠	٢٠٠٣	٨٤٠٧
١٩٥٠	١٦١٦	٧٣٣٧	١٩٦٢	٢٠١٧	٨٦٣٩
١٩٥٢	١٥٧٦	٧٢٧٣	١٩٦٤	١٩٩٠	٨٥٣١
١٩٥٤	١٧١٠	٧٤٢٩	١٩٦٦	٢٠٩١	٨٣٧٧
١٩٥٦	١٧٣٦	٧٥٢٥			

١– نختار سنة ولتكن سنة ١٩٤٦ ونعطيها رقماً قياسياً مقداره ١٠٠ (ويمكن اختيار أى سنة من سنوات السلسلة لتكون سنة الأساس، وقد يؤخذ متوسط عدد السنين كسنة الأساس).

٢– تحول أعداد كل من الماشية والأغنام منسوب إلى سنة الأساس. فمثلاً عدد الأغنام فى عام ١٩٤٨ يساوى ٦٧٣١٠٠٠، وعددها فى عام ١٩٤٦ هو ٦٣٥٤٠٠٠، ونسبة عدد الأغنام فى عام ١٩٤٨ إلى عددها ١٩٤٦ هو:

$$\frac{٦٧٣١٠٠٠}{٦٢٥٤٠٠٠} \times ١٣٠٠ - ٪٩٨$$

٣– بنفس الطريقة السابقة يمكن الحصول على الأرقام القياسية لأعداد كل من الماشية والأغنام باعتبار سنة ١٩٤٦ هى سنة الأساس لكل من الماشية والأغنام كما فى الجدول رقم (١٣–٢).

جدول رقم (١٣-٢)

الأرقام القياسية لتطور أعداد الماشية والأغنام فى اسكتلندا

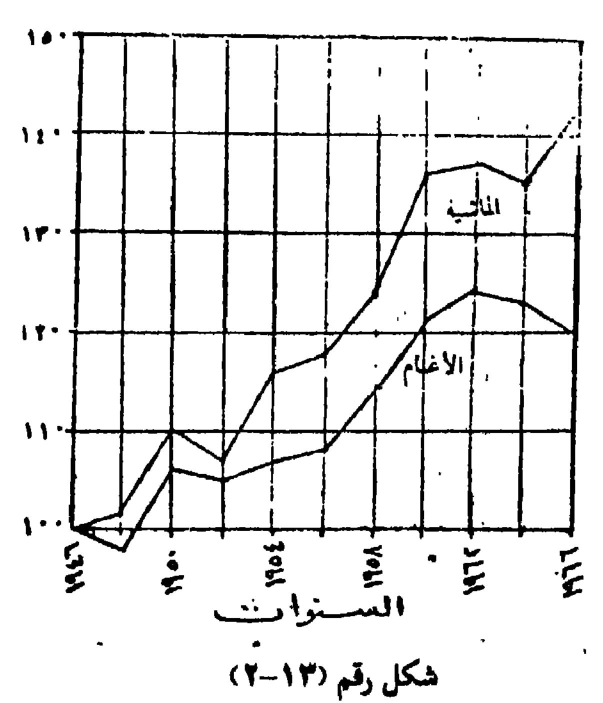
السنة	الأرقام القياسية		السنة	الأرقام القياسية	
	عدد الماشية	عدد الأغنام		عدد الماشية	عدد الأغنام
١٩٤٦	١٠٠	١٠٠	١٩٥٨	١٢٤	١١٤
١٩٤٨	١٠٢	٩٨	١٩٦٠	١٣٦	٢١
١٩٥٠	١١٠	١٠٦	١٩٦٢	١٣٧	١٢٤
١٩٥٢	١٠٧	١٠٥	١٩٦٤	١٣٥	١٢٣
١٩٥٤	١١٦	١٠٧	١٩٦٦	١٤٢	١٢٠
١٩٥٦	١١٨	١٠٨			

٤- تقوم بتمثيل الأرقام القياسية بيانياً بطريقة الخطوط البيانية كما في الشكل رقم (١٦-٢) والذى يمثل الأرقام القياسية لتطور أعداد الماشية والأغنام من عام ١٩٤٦ حتى ١٩٦٦.

وتتميز طريقة الأرقام القياسية بسهولة قراءة وملاحظة التغير النسبى للظاهرة خلال فترة الدراسة. فمثلاً الرقم القياسى ١٣٦ يعطينا دلالة على نمو وتطور الظهرة مقدارة ٣٦٪ عن سنة الأساس، والرقم القياسى ٩٨ يعنى نقص أو اضمحلال الظاهرة موضع الدراسة بما يساوى ٢٪ عن سنة الأساس.

ومن مميزات هذه الطريقة أيضاً أنها تلغى القيم الأصلية وتحولها إلى نسب مئوية من سنة الأساس مما يسهل عملية المقارنة ودراسة أوجه الشبه والاختلاف بين جميع القيم للظواهر موضع التحليل. فمثلاً من الشكل رقم (١٣-٢) يمكن أن نوضح أن سرعة تزايد أعداد الماشية خلال فترة السلسلة الزمنية أكبر من سرعة تزايد أعداد الأغنام. وذلك على عكس مايظهر عند دراسة الرسم البيانى الحسابى أو العددى (شكل رقم ١٣-١) مما يؤكد أن الرسوم البيانية التى توضح التغيرات العددية للقيم الأصلية للظاهرة قد تكون مضللة أحياناً. أما أهم عيوب طريقة الأرقام القياسية كوسيلة لبيان تطور الظاهرة أنها لا توضح إلا التغير النسبى بين سنوات السلطة

الزمنية وسنة الأساس فقط. وعلى هذا فإن التغير بين سنوات السلسلة نضمها لايمكن معرفته إلا إذا كان منسوباً إلى سنة الأساس. فمثلاً فمن الشكل رقم (١٦-٢) نجد أن الأرقام القياسية للماشية والأغنام قد ارتفع بين سنة ١٩٥٦ وسنة ١٩٥٨ بمقدار ٦ (من ١١٨ - ١٢٤ للماشية، ومن ١٠٨ - ١١٤ للأغنام) غير أن انحدار الخط البيانى على الرسم للظاهرتين متشابه بين هاتين السنتين. والحقيقة أن هناك اختلافا واضحاً بين هاتين القيمتين لكل من الظاهرتين، حيث أن ٦ أرقام قياسية من أصل ١١٨ لاتمثل إلا زيادة قدرها ٥ ٪ فى عدد الماشية، فى حين أن ٦ أرقام قياسية من أصل ١٠٨ تمثل زيادة قدرها ٦ ٪ تقريباً فى عدد الأغنام. وجدير بالذكر أن اختيار معدل عدد من السنوات (خمس سنوات متتالية مثلاً) كسنة أساس يكون أفضل من اختيار سنة بذاتها لتكون سنة الأساس وذلك لتجنب الذبذبات Fluctiuations التى قد تطرأ على تطور ونمو الظاهرة بسبب العوامل المتحكمة والمؤثرة فى هذا التطور والنمو.



شكل رقم (١٣-٢)

الأرقام القياسية لتطور أعداد الماشية والأغنام فى اسكتلندا

وهناك طريقة أخرى لبيان التغير النسبي قد يطرأ على ظاهرة معينة خلال فترة زمنية محددة هي طريقة الرسم البياني اللوغاريتمي والنصف لوغاريتمي الذى قمنا بشرح فكرة رسمه وانشائه فى الفصل الثالث من هذا الكتاب. ويسود استخدام طريقة الرسم البياني النصف لوغاريتمي فى تمثيل معدلات التغير أو النمو للظاهرة والتى تتغير تغيراً زمنياً مثل ظاهرة نمو السكان. فإذا أريد تمثيل التيغر الذى يطرأ على مثل هذه الظاهرة يجب استخدام الرسم البيانى النصف لوغريتمي والذى يمثل فيه المحور الأفقى عامل الزمن بالطريقة الحسابية، بينما يمثل المحور الرأسى مقدار تزايد الظاهرة لوغريتمياً. فمثلاً أهم ما يلاحظ على طريقة إنشاء الشكل رقم (١٣-٣) الذى يوضح النمو السكانى لمدينتين أ، ب لسنوات التعداد من ١٨٠١ إلى ١٩٦١ (جدول رقم ١٣-٣) أن تقسيم المحور الأفقى يتفق مع عدد سنوات التعدادات بينما نجد على المحور الرأسى تقسيماً لوغاريتمياً أى أنه مقسم إلى ثلاث من الدورات اللوغاريتمية تبدأ من ١٠٠٠ إلى ١٠٠٠٠، والثانية من ١٠٠٠٠ إلى ١٠٠٠٠٠، والثالثة من ١٠٠٠٠٠ إلى ١٠٠٠٠٠٠ ولكنها لم تستكمل وانتهت عند رقم ٣٠٠٠٠٠، وبنظرة أخرى إلى الشكل نجد على جانبه الايمن دورتين لوغاريتميتين الأولى تبدأ من ١٠٠ إلى ١٠٠٠ والثانية تبدأ من ١٠٠٠ إلى ١٠٠٠٠. وأهم مميزات الرسم البيانى النصف لوغاريتمى السابق أن المقياس على محوريه الرأسيين (الأيسر والأيمن) يساعد على سهولة المقارنة بين ظاهرتين إحداهما مقدارها كبير جداً والأخرى مقدارها صغير جداً. ففى الشكل نجد خطين بيانيين نصف لوغاريتميين يوضحان نمو السكان وتطور عددهم فى مدينتين إحداهما ازداد عدد سكانها من ٢٨٨٠١ نسمة إلى ٣١١٨٩٩ نسمة، بينما الأخرى ازداد عدد سكانها من ٨٦٨ نسمة إلى ٥١٨٥ نسمة. وهذا النمط من التطور لايمكن بيانه بغير الخط البيانى اللوغاريتمى أو النصف لوغاريتمي. كما يفيدنا هذا النوع من الرسوم البيانية فى معرفة الفترة الزمنية التى تتضاعف فيها قيمة الظاهرة، وذلك بسبب ماتتميز به الأشكال اللوغاريتمية من أن المسافة على المحور الرأسى بين الرقمين ١٠٠٠، ٢٠٠٠ تساوى المسافة التى تقع بين الرقمين ٣٠٠٠، ٦٠٠٠، وتساوى أيضاً المسافة بين الرقمين ٤٠٠٠٠، ٨٠٠٠٠ وكذلك بين الرقمين ٣٠٠٠٠، ٦٠٠٠٠. وبعبارة أخرى أن المسافات على المحو الرأسى

الذى يوضح المقياس اللوغاريتمى تكون متساوية بين أى رقمين النسبة بينهما كنسبة ١ : ٢ . فمثلاً إذا أردنا معرفة الفترة الزمنية التى يتضاعف فيها السكان فى المدينة «أ» فى الشكل رقم (١٣-٣)، فإننا نأخذ نقطتين على الرسم احداهما تمثل عدد سكان المدينة فى فترة ما، ثم نحدد النقطة الأخرى على الخط البيانى نفسه عندما يتضاعف عدد السكان. فنقطة س تمثل عدد سكان المدينة الذى يبلغ ٣٠٠٠٠ نسمة، ونقطة ص تمثل عدد سكان المدينة حينما بلغ ٦٠٠٠٠ نسمة، فإذا أسقطنا عمودين من هاتين النقطتين على المحور الأفقى نجد أن المسافة بين مسقطيهما تقع بين عامى ١٨٠٦ ، ١٨٥٣ . وبذلك تستنتج أن عدد السكان فى المدينة «أ» قد تضاعف فى النصف الأول من القرن التاسع عشر خلال فترة ٤٧ سنة تقريباً.

جدول رقم (٦-٣)

تعداد السكان لمدينتين أ، ب

١٨٠١ - ١٩٦١

السنة	تعداد السكان		السنة	تعداد السكان	
	المدينة أ	المدينة ب		المدينة أ	المدينة ب
١٨٠١	٢٨٨٠١	١٨٦٨	١٩٠١	٢٣٩٧٤٣	٢٤٩٣
١٨٢١	٤٠١٩٠	١١٣٨	١٩٢١	٢٦٢٦٢٤	٢٨٧٧
١٨٤١	٥٢١٦٤	١٨٣٥	١٩٤١	٢٧٦١٨٩	٣٠٦٤
١٨٦١	٧٤٦٩٣	٢٢٨٣	١٩٦١	٣١١٨٩٩	٥٠٨٥
١٨٨١	٨٦٥٧٥	٢٦٣٨			



السنوات

شكل رقم (١٣-٣)

تمثيل النمو السكانى لمدينتين أ، ب على الرسم البيانى النصف لوغاريتمى

التغيرات (التحركات) المميزة فى السلاسل الزمنية:

يمكننا الآن القول - بعد أن عرفنا مميزات الرسم البيانى للسلسلة الزمنية أن تطور أو نمو الظاهرة الجغرافية قيد التمثيل هو بمثابة نقطة تتحرك مع مرور الزمن تحت تأثير قوى معينة (طبيعية، بشرية، اقتصادية ..الخ) وذلك فيما يشبه لتحرك

المادى للذرة تحت تأثير قوى مادية مختلفة. ولكن من دراسة وملاحظة كثير من السلاسل الزمنية يمكن اكتشاف وجود تغيرات (تحركات) Changes أو اختلافات مميزة وبدرجات مختلفة.. وجدير بالذكر أن تحليل مثل هذه لاتغيرات له أهمية كبرى فى التنبؤ بالتغيرات المستقبلية الظاهرة قيد الدراسة. وبناء على ذلك يمكن تصنيف التغيرات فى السلاسل الزمنية إلى أربعة أنماط تسمى غالباً «مكونات السلسلة الزمنية» نوجزها فيمايلى:

١- التغيرات طويلة المدى (الاتجاه العام) The Trend. تشير هذه التغيرات إلى الاتجاه العام (الحركة العامة) الذى يكون عليه الشكل البيانى للسلسلة الزمنية على مدار فترة طويلة من الزمن. ويرمز للاتجاه العام على الرسم البيانى للسلسلة الزمنية بمنحنى أو خط الاتجاه العام الذى يمثل التدرج والانتظام والاستمرار فى السلسلة الزمنية.

٢- التغيرات الدورية Periodioal (Cyclical) Variations. وتشير هذه التغيرات إلى الذبذبات طويلة المدى حول خط الاتجاه العام. وتسمى هذه التغيرات أحياناً «دورات» التى قد تتبع أولاً تتبع نفس النمط بعد كل فترة زمنية متساوية. وفى مجال الجغرافية الاقتصادية تعتبر التغيرات دورية إذا تكررت بعد فترات زمنية تزيد عن سنة.

٣- التغيرات الموسمية Seasonal Variation تمثل هذه التغيرات النمط المتماثل، أو المنتظم لحركة السلسلة الزمنية فى الأشهر المتقابلة خلال السنوات المتتالية، وتتمثل هذه التغيرات فى لحادثات التى تقع شهرياً أو ربع سنوياً أو سنوياً مثل زيادة مبيعات المحلات التجارية فى الفترة السابقة للأعياد أو موسم الحج، زيادة كمية الأمطار خلال فصل الشتاء عنه أثناء فصل الصيف فى الأقاليم شبه الجافة فى العروض المعتدلة فى غرب القارات. وعلى الرغم من أن التغيرات الموسمية تشير إلى الدورة السنوية، فإنها يمكن أن تمتد لتشمل الدورية لأية فترة من الزمن مثل ساعة، يوم، أسبوع ... الخ اعتماداً على نوع البيانات المتاحة. وأهم أسباب التغيرات الموسمية فى السلسلة الزمنية هى العادات والتقاليد الإجتماعية والعوامل المناخية.

٤ - التغيرات العشوائية Random Variations. وتشير هذه التغيرات إلى الحركة غير المنتظمة فى السلسلة الزمنية، بمعنى أنه ليس لها نمط معين أو قاعدة ثابتة، أى أنها تنجم فى الغالب عن عوامل عارضة أو فجائية لايمكن التنبؤ بمواعيد حدوثها بدقة مثل الفيضانات والزلازل ونتائج الانتخابات ... وغيرها. وعلى الرغم أنه من المعتاد افتراض أن مثل هذه الحادثات تنتج تغيرات تستمر لفترة قصيرة من الزمن فمن المعقول أن تكون على درجة من الكثافة نتيجة لوجود دورات جديدة أو غيرها من التغيرات ولابد من استبعادها عند قياس الاتجاه العام أو عند قياس التيغرات الموسمية للظاهرة موضع التحليل.

وعموماً فإن الظاهرة الواحدة قد تتأثر بجميع هذه التغيراتُ أو بعضها الأمر الذى ينجم عنه تقلبات صاعدة أو هابطة فى خط سير الظاهرة. ولهذا السبب يهتم الجغرافيون بدراسة مثل هذه التغيرات ومعرفة أسبابها لتفادى آثارها الضارة فى المستقبل.

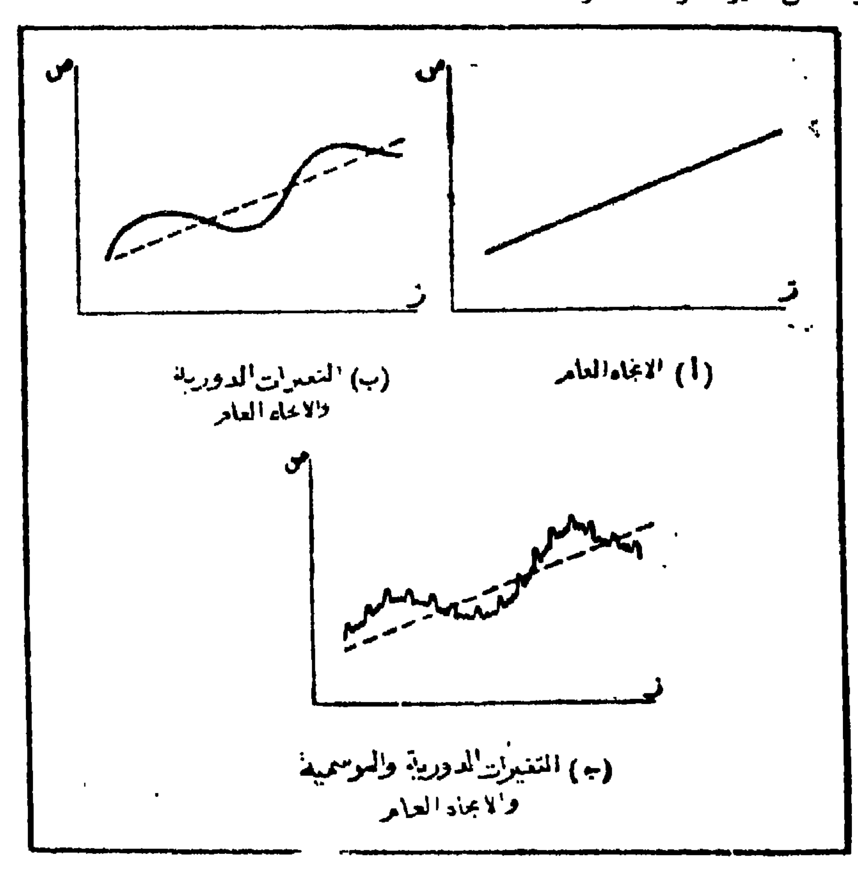
تحليل السلسلة الزمنية:

يهدف تحليل السلسلة الزمنية إلى إيجاد طريقة مناسبة لقياس التغيرات وبالتالى دراسة علاقاتها بالظروف المختلفة. ويتضمن تحليل السلاسل الزمنية وصف مكونات التغيرات الموجودة داخل السلسلة، أو بعبارة أخرى إيجاد مركبات القوى التى تؤثر فى خط سير الظاهرة خلال الفترة الزمنية موضع التحليل. ولتوضيح الطرائق التى تستخدم فى هذا الوصف نشير إلى الشكل رقم (١٣-٤) والذى يعرف بالسلسلة الزمنية المثالية التى تتخذ فيها المركبات (القوى) أشكالاً بيانية مختلفة.

ويمكن أن يعطينا عرض الأشكال البيانية السابقة (شكل رقم ١٣-٤أ، ب، جـ) أسلوباً خاصاً لتحليل السلاسل الزمنية. فتفترض أن المتغير (ص) الذى يعبر عن السلسلة الزمنية هو حاصل ضرب المتغيرات ن، د، و، ش والتى ينتج عنها الاتجاه العام (ت) والتغيرت الدورية (د) والتغيرات الموسمية (و) والتغيرات العشوائية (ش). أى أن قيمة الظاهرة تساوى حاصل ضرب مركباتها الأساسية. وباستخدام الرموز نجد أن:

ص = ت × د × و × س = ت د و س

ومن هذا المنطلق فإن تحليل السلاسل الزمنية يتضمن فحص المتغيرات ت، د، و، ش السابقة والتي يشار إليها بمفكوك السلسلة الزمنية أو المكونات الأساسية لمتغيراتها. وتجدر الإشارة إلى أن بعض الاحصائيين يفضلون اعتبار قيمة الظاهرة مساوية لمجموع قيم مركباتها، أو بعبارة أخرى ص = ت + د + و +س. وكل من الطريقتين السابقتين لتحديد قيمة الظاهرة يمكن استخدامها فى تحليل السلسلة الزمنية. ولكى نعين تأثير كل قوة من قوى التأثير (قوة الاتجاه العام، قوة التغيرات الدورية، الموسمية العشوائية) على السلسلة الزمنية يجب عزل كل قوة من هذه القوى عن تأثير القوى الأخرى.



شكل رقم (١٣-٤) مكونات السلسلة الزمنية المثالية

أولاً: تقدير الاتجاه العام:

يتناول الجغرافيون دراسة الظواهر التى تتزايد بطبيعتها خلال فترة زمنية مثل كميات الإنتاج للمحاصيل أو السلع الصناعية والمعدنية، وعدد السكان وحجم الاستهلاك، أو الظواهر التى تتناقص على مدار الزمن مثل معدل الوفيات واستهلاك السلع. ويكون أحسن تمثيل بياني للإتجاه العام للسلاسل الزمنية لمثل هذه الظواهر أن نمهد لها بطريقة مناسبة، خطاً مستقيماً أو منحنى أملسا يمثل سير الظاهرة لو لم تؤثر عليها عوامل أخرى تؤدى إلى تفاوت فى مقدار الظاهرة. وعندما يتم رسم خط – أو تمهيد منحنى – الاتجاه العام فإننا نجد أن كل فترة زمنية يناظرها قيمتان للظاهرة إحداهما هى «القيمة المشاهدة الحيقيقة والأخرى هى القيمة الاتجاهية، وهى القيمة التى يحددها الخط أو المنحنى الممهد ، وهي بطبيعة الحال تختلف عن القيمة المشاهدة. وتجدر الإشارة إلى أن متوسط القيم الاتجاهية يساوى متوسط القيم المشاهدة. ويمكن معرفة وتقدير الاتجاه العام Trend للظاهرة بعدة طرق هى: طريقة خط (منحنى) الاتجاه العام وتمهيده باليد، طريقة أنصاف المتوسطات (أشباه المتوسطات) ، طريقة المتوسطات المتحركة، طريقة المربعات الصغرى. وسنلقى الضوء فيمايلى على كل طريقة على حدة لإبراز أهم مميزاتها وعيوبها عند اتخاذها كأساس لتحليل السلسلة الزمنية.

١ – طريقة التمهيد باليد: تعتمد هذه الطريقة على توفيق أو تمهيد خط الاتجاه العام باليد بحيث يمر بأكبر عدد ممكن من النقط – التى تمثل قيم الظاهرة فى كل وحدة زمنية على الرسم البيانى– أو يمر بين النقط باتزان وذلك بعد رسم المنحنى التاريخى للسلسلة عن طريق توصيل النقط بخطوط منكسرة. ويمكن أن نستنتج من الرسم البيانى بصورة عامة شكل الاتجاه العام، بمعنى هل هو خط مستقيم فيكون له معادلة من الدرجة الأولى ، أو منحنى من نوع معين قد تكون معادلته من الثانية أو الثالثة أو من درجة أعلى من ذلك.

ويمثل الشكل رقم (١٣– ٥) منحنياً تاريخياً وخطاً ممهداً باليد للسلسلة الزمنية التى توضح تطور إجمالى الصادرات لجمهورية مصر فى الفترة ١٩٦٠ – ١٩٧٢ (جدول رقم ١٣–٤).

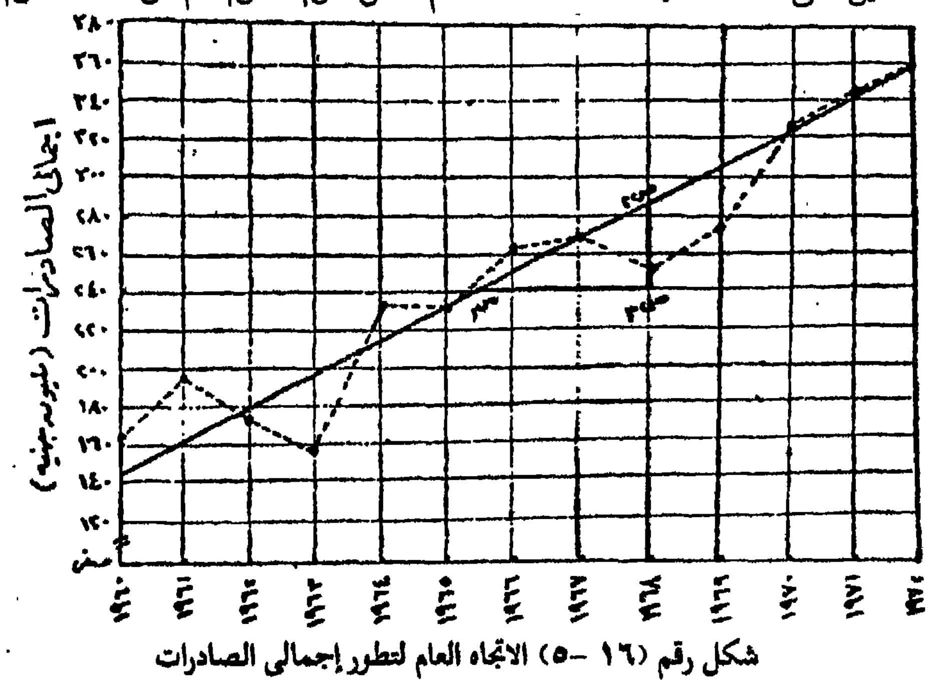
جدول رقم (١٣-٤)

تطور إجمالي الصادرات لجمهورية مصر (بالألف جنيه مصري)

في الفترة ١٩٦٠ - ١٩٧٢

السنة	الصادرات	السنة	الصادرات	السنة	الصادرات
١٩٦٠	١٦٠٤٥٥	١٩٦٥	٢٣٤٣٧٧	١٩٧٠	٣٢٣٩٢٩
١٩٦١	١٩٧٧٥٧	١٩٦٦	٢٦٣١٣٣	١٩٧١	٣٤٣١٧٧
١٩٦٢	١٦٨٩٢٦	١٩٦٧	٢٦٣١٣٥	١٩٧٢	٣٥٨٧٧٥
١٩٦٣	١٥٨٥١٠	١٩٦٨	٢٤٦١٣٧		
١٩٦٤	٢٣٦٧٩٨	١٩٦٩	٢٧٠٢٩٥		

ومن الخط الممهد يمكن حساب مقدار التغير الذي يطرأ على الظاهرة في كل وحدة زمنية في المتوسط - وهو مايطلق عليه «معدل التغير» - وذلك بأن نحدد نقطتين على الخط الممهد (خط الاتجاه العام) مثل ص١، ص٢ ثم من النقطة ص٢



شكل رقم (١٦ -٥) الاتجاه العام لتطور إجمالي الصادرات

لجمهورية مصر في الفترة ١٩٦٠ - ١٩٨٠

نرسم عموداً يوازى المحور الرأسى، ومن النقطة ص١ نرسم خطا يوازى المحور الأفقى بحيث يلتقى الخطان عند النقطة ص٣ فيكون معدل التغير هو ميل الخط الممهد (أى ظل الزاوية صَ١) وهو يساوى الاحداثى الرأسى ص٢، ص٣ بوحدات المحور الرأسى (إجمالى الصادرات بالمليون) مقسوماً على الاحداثى الأفقى ص١، ص٣ بوحدات المحور الأفقى (الوحدات الزمنية). ومن الشكل رقم (١٣-٥) يتبين أن معدل إجمالى الصادرات فى المتوسط ١٨,٥ مليون جنيه. وعلى فرض أن الاتجاه العام يأخذ شكل الخط المستقيم الذى معادلته ص = م س ± حـ، فإن معدل التغير (م)، وهو ميل الخط، يكون ثابت على جميع أجزاء الخط مهما أختلفت الفترة الزمنية المحسوب عندها المعدل. وبذلك يمكن تحديد المعادلة السالفة بمعلومية معدل التغير (م) والجزء المقطوع من المحور الرأسى (حـ) والذى يمثل هنا الحد الأدنى لإجمالى الصادرات سنوياً (يمكن تحديده من الرسم بالمقدار ١٢٥ مليون جنيه). وعلى ذلك فإن معادلة الخط المستقيم لهذه السلسلة تكون:

ص = ١٨,٥ س + ١٢٥

ويعاب على طريقة تعيين الاتجاه العام التمهيد باليد بأنها ليست دقيقة إذ أن العمل فيها يتم بطريقة تقديرية تختلف من شخص إلى آخر، كما تتوقف على مدى خبرة ومران الباحث فى تمهيد خط مستقيم باليد ليمر بأكبر عدد ممكن من النقط. وعلى الرغم من ذلك فإن خط الاتجاه العام الممهد بالبد يمكن استخدامه كمؤشر للتنبؤ بقيم الظاهرة فى فترة لاحقة لفترة السلسلة الزمنية موضع التحليل، إذ أن أحداثى أية نقطة تقع على خط الاتجاه عبارة عن إجمالى الصادرات عند وحدة زمنية مناظرة. فعلى سبيل المثال يمكن حساب القيمة المقدرة لإجمالى الصادرات لعام ١٩٨٥ كالآتى:

القيمة الاتجاهية (ص) = ١٨,٥ س + ١٢٥

قيمة س (فى هذه الحالة على أساس أن نقطة الأصل هى ١٩٦٠)

= ١٩٨٥ - ١٩٦٠ = ٢٥ بالتعويض فى معادلة خط الاتجاه العام نجد:

ص = ١٨,٥ × ٢٥ + ١٢٥ = ٥٨٧,٥ مليون جنيه

٢- طريقة أنصاف (أشباه) المتوسطات Semi-Average: وهى طريقة تقريبية أيضاً لتعيين خط الاتجاه العام ولكنها أفضل من التمهيد باليد. ولتوفيق خط مستقيم يوضح الاتجاه العام لسير الظاهرة تقسم البيانات إلى قسمين متساويين - كلما أمكن - ثم نحصل على المتوسط الحسابى لكل قسم على حدة، وهذا يعطينا قيمة نقطتين على خط السلسلة الزمنية (فى حالة إذا كانت السلسلة تحتوى على عدد زوجى من السنين فإنه يمكن تقسيمها إلى قسمين متساويين بإهمال السنة التى تقع فى منتصف السلسلة). ولرسم خط الاتجاه العام نعين على الرسم النقطتان التى تمثل كل منهما قيمة كل متوسط أمام منتصف الفترة الزمنية لكل مجموعة ثم نصل بين هاتين النقطتين، وبذلك يمكن تحديد القيم الاتجاهية.

مثال (٣):

تبين البيانات فى الجدول رقم (١٣-٥) الإنتاج السنوى من الفحم (مليون كيلو جرام) فى دولة مابين سنة ١٩٤٨ وسنة ١٩٥٨ ولتحليلها بطريقة أنصاف المتوسطات بهدف تحديد الاتجاه العام للإنتاج خلال هذه الفترة نجرى الخطوات التالية:

جدول رقم (١٣-٥)

الإنتاج السنوى من الفحم (مليون كيلو جرام)

فى دولة ما الفترة ١٩٤٨ - ١٩٥٨

السنة	١٩٤٨	١٩٤٩	١٩٥٠	١٩٥١	١٩٥٢	١٩٥٣	١٩٥٤	١٩٥٥	١٩٥٦	١٩٥٧	١٩٥٨
الإنتاج	٥٠,٠	٣٦,٥	٤٣,٠	٤٤,٥	٣٨,٩	٣٨,١	٣٢,٦	٣٨,٧	٤١,٧	٤١,١	٣٣,٨

١- تحتوى هذه السلسلة الزمنية على بيانات ١١ سنة فنقسمها إلى قسمين متساويين كل نهما ٥ سنوات (مع حذف السنة فى منتصف السلسلة وهى سنة ١٩٥٣)، فيحتوى القسم الأول على السنوات ١٩٤٨ - ١٩٥٢، والقسم الثانى على السنوات ١٩٥٤ - ١٩٥٨.

٢- مجموع الإنتاج السنوى للقسم الأول يساوى ٢١٢,٩ ، ومجموع الإنتاج السنوى للقسم الثاني يساوى ١٨٧,٩ .

٣- المتوسط الحسابى للقسم الأول = ٢١٢,٩ ÷ ٥ = ٤٢,٦ ، والمتوسط الحسابى للقسم الثانى = ١٨٧,٩ ÷ ٥ = ٣٧,٦ . توقع القيمتان على الرسم البيانى (شكل رقم ١٣-٦) مقابل السنة الوسطى فى القسم الأول وهى سنة ١٩٥٠ ، ومقابل السنة الوسطى فى القسم الثانى وهى سنة ١٩٥٦ (سَ، سً على الترتيب) ، فإذا وصلنا بينهما فإن الخط الذى يمر بينهما هو خط الاتجاه العام للظاهرة، والنقط على هذا الخط التى تقابل السنين هى القيم الاتجاهية لهذه السنين.

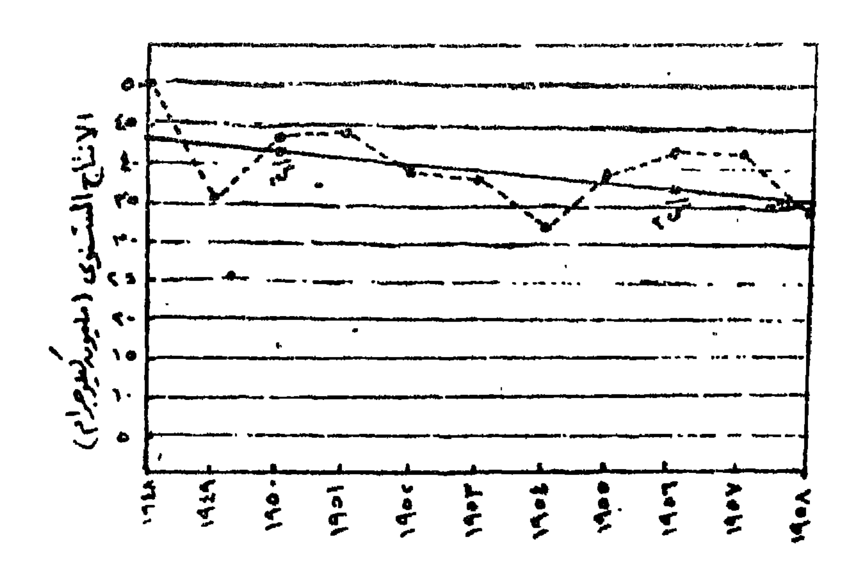
٤- تعد القيمة الاتجاهية لسنة ١٩٥٠ هى ٤٢,٦ مليون جرام، والقيمة الاتجاهية لسنة ١٩٥٦ هى ٣٧,٦ مليون جرام، فإذا كانت النتائج التى حصلنا عليها تدل على أنه فى ٦ سنوات (١٩٥٠ - ١٩٥٦) حدث إنخفاض يساوى ٥ مليون كيلو جرام (٤٢,٦ - ٣٧,٦) أى بانخفاض نسوى مقداره ٥,٠ ÷ ٦ = ٠,٨٣ مليون كيلو جرام/ سنة، وهو ما أطلقنا عليه سلفا معدل التغير (ميل الخط المستقيم). وبعبارة أخرى إذا كان خط الاتجاه العام خطاً مستقيماً فإن ميله (م) يساوى الفرق بين المتوسطين الحسابيين لمنتصفى السلسلة الزمنية مقسوماً على الفرق بين زمنيهما، أى أن:

$$م = \frac{٤٢,٦ - ٣٧,٦}{١٩٥٦ - ١٩٥٠} = \frac{٥,٠}{٦} = ٠,٨٣$$

٥- إذا كانت معادلة خط الاتجاه العام هى ص = م س ± حـ، وإذا أخذنا نقطة عند سنة ١٩٥٠ فإن (حـ) تساوى المتوسط الحسابى للقسم الأول وهو ٤٢,٦ وتكون معادلة الاتجاه العام هى: ص = ٠,٨٣ س + ٤٢,٦ .

٦- لتعيين القيمة الاتجاهية لأية سنة يجب أن نأخذ فى الحسبان السنة التى أخذنا عندها نقطة الأصل فمثلاً سنة ١٩٥١ تقابل س = -١ ، فتكون القيم الاتجاهية لسنة ١٩٥١ هى:

$$ص = ٠,٨٣ \times (-١) + ٤٢,٦ = -٠,٨٣ + ٤٢,٦ = ٤١,٧٧$$



شكل رقم (١٣-٣)

المنحنى التاريخى وخط الاتجاه العام بطريقة أنصاف المتوسطات

للإنتاج السنوى من الفحم فى دولة ما للفترة ٤٨-١٩٥٨

وإذا أخذنا نقطة الأصل عند سنة ١٩٥٦ فإن (حـ) تساوى المتوسط الحسابى للقسم الثانى (٣٧,٦) وتكون معادلة الاتجاه العام هى : ص = ٠,٨٣ س + ٣٧,٦ وتكون القيمة الاتجاهية لسنة ١٩٥٨ (التى تقابل س = −٢) كالآتى:

ص = ٠,٨٣ × (−٢) + ٣٧,٦ = −١,٦٦ + ٣٧,٦ = ٣٥,٩٤

٧- من الشكل البيانى (شكل رقم ١٣-٦) يتضح أن هناك إتجاهاً عاماً هابطاً Downward Secular Trend ولكنه هبوط بسيط جداً، أى أن الإنتاج السنوى إنخفض بشكل ضئيل فى غضون الفترة الزمنية موضع الدراسة. ومن الواضح كذلك أن اختلافات القيم الأصلية وانحرافها عن خط الاتجاه العام يمكن تلخيصها بالإشارات الموجبة والسالبة الآتية: + − + + − − − + + + −.
وإذ كان الانحراف عن خط الاتجاه العام شديداً فإنه يجب علينا أن نحدد طبيعته، بمعنى هل هذا التفاوت والانحراف عن خط الاتجاه العام عشوائياً؟ أم أنه يتبع نظاماً محدداً أو نمطاً معيناً؟

وعلى الرغم من أن هذه الطريقة بسيطة فى تطبيقها، إلا أنها قد تؤدى إلى نتائج غير دقيقة إذا استخدمت بدون تمييز. فكما عرفنا أن حساب ميل خط الاتجاه لعام يتوقف على حساب كل من المتوسطين الحسابين ولا يخفى علينا أن حسابهما يتأثر بالقيم المتطرفة أو الشاذة (الشديدة الارتفاع أو الشديدة الانخفاض أو الاثنين معاً)، فإذا كان أحد القسمين يحتوى على قيم مرتفعة أو منخفضة أكثر من القسم الآخر فإن ذلك يؤثر على قيمتى المتوسطين الحسابيين وبالتالى لا يكون خط الاتجاه العام المحسوب ممثلاً الحالة بدقة كاملة. ومن عيوب هذه الطريقة أيضاً أنها قابلة للتطبيق فقط فى حالة إذا كان الاتجاه خطاً أو يقرب إلى خطين، على الرغم من أنه يمكن مد صلاحيتها فى الحالات التى يمكن تقسيم البيانات فيها إلى عدد من الأقسام فى كل قسم يكون الاتجاه العام فيه خطياً.

٣- طريقة المتوسطات المتحركة The Moving Averages: عرفنا الغرض الأساسى من تعيين خط الاتجاه العام هو تقدير قيم اتجاهية لاتقع إلا تحت تأثير عامل واحد فقط هو أثر الاتجاه العام، أى تقليل أثر المتغيرات الدورية، الموسمية والعشوائية. وقد وجدنا أن تعيين الاتجاه بالتمهيد باليد وبطريقة أنصاف المتوسطات ليست بالدقة الكافية التى يعتمد عليها فى التنبؤ. وحفاظاً على الهدف الرئيسى من الاتجاه العام للظاهرة يمكن تعيين الاتجاه العام بأسلوب أسهل باستخدام مايسمى «بالمتوسطات المتحركة»، وهى عبارة عن استخراج متوسط قيم عدد معين (ثلاث أو خمس سنوات مثلا) من السنوات المتعاقبة أو المتداخله (فترة فى السلسلة الزمنية)، ثم تثبت قيم هذه المتوسطات أمام السنوات الوسطى لكل فترة فى السلسلة الزمنية. ويمكن تحديد طول الفترة فى السلسلة الزمنية من ملاحظة المنحنى التاريخى وذلك عن طريق معرفة الفرق الزمنى بين كل قيمتين متتاليتين أو كل قاعدتين متتاليتين.

فمثلاً إذا فرضنا أن طول كل فترة فى السلسلة (ن) يساوى ٣ سنوات وأن قيم السلسلة الزمنية هى : $س_1$، $س_2$، $س_3$، $س_4$ .. $س_م$، فإن المتوسطات المتحركة - أو القيم الاتجاهية - من الدرجة ن يمكن أن نحصل عليها بمتتابعة من المتوسطات الحسابية:

(أ) القيمـة الاتجاهيـة الأولى = $\frac{\text{س}_{١} + \text{س}_{٢} + \text{س}_{٣}}{٣}$ وتمثل قيمة اتجاهية للسنة $\text{س}_{٢}$

(ب) القيمة الاتجاهيـة الثانية = $\frac{\text{س}_{٢} + \text{س}_{٣} + \text{س}_{٤}}{٣}$ وتمثل قيمة اتجاهية للسنة $\text{س}_{٣}$

(حـ) القيمة الاتجاهية الثالثة = $\frac{\text{س}_{٣} + \text{س}_{٤} + \text{س}_{٥}}{٣}$ وتمثل قيمة اتجاهية للسنة $\text{س}_{٤}$

(د) القيمة الاتجاهية الأولى = $\frac{\text{س}_{م-٢} + \text{س}_{م-١} + \text{س}_{م}}{٣}$ وتمثل القيمةالاتجاهية للسنة قبل الأخيرة والتى ترتيبها م -١

وتسميى المجاميع $\text{س}_{١} + \text{س}_{٢} + \text{س}_{٣}$، $\text{س}_{٢} + \text{س}_{٣} + \text{س}_{٤}$ ، + $\text{س}_{٣}$ $\text{س}_{٤}$ $\text{س}_{٥}$ ، $\text{س}_{م-٢} + \text{س}_{م-٤} + \text{س}_{م}$ بالمجاميع المتحركة من الدرجة ن.

أما إذا كانت (ن) زوجية، أى ٤ سنوات مثلاً، فإن:

(أ) القيمة الاتجاهية الأولى = $\frac{\text{س}_{١} + \text{س}_{٢} + \text{س}_{٣} + \text{س}_{٤}}{٤}$ وتمثل القيمـة الاتجاهية غير المركزية أى لمنتصف السنة الثانية ($\text{س}_{٢}$) ويرمزلها بالرمز $\text{ل}_{١}$

(ب) القيمـة الاتجاهية الثانية = $\frac{\text{س}_{٢} + \text{س}_{٣} + \text{س}_{٤} + \text{س}_{٥}}{٤}$ وتمثل القيمـة الاتجاهية غير المركزية أى لمنتصف السنة الثالثة ($\text{س}_{٣}$) ويرمز لها بالرمز $\text{ل}_{٣}$.

وحتى نحصل على القيم الاتجاهية لسنوات السلسلة نفسها بدلاً من منتصف سنوات السلسلة فإننا نحسب القيم الاتجاهية على أساس أنها متوسط القيم لكل فترتين متتاليتين. وتكون القيمـة الاتجاهية للسنة الثالثة فى السلسلة $\frac{\text{ل}_{١} + \text{ل}_{٢}}{٢}$

ويمكن التوصل إلى نفس النتيجة بقسمة المجموع المتحرك لكل فترتين طول كل منهمـا ٤ سنوات على مجمـوع سنوات الفـتـرتين وهى ٨ سنوات. وهكذا فإننا نحصل على متوسطات حسابية تتحرك أمام كل سنة من سنوات السلسلة الزمنية

وتمثل قيمها الإتجاهية. ومن المعتاد أ نضع قيمة المتوسطات المتحركة فى مكانها الملائم بالنسبة للبيانات الأصلية كما فى الشكل رقم (١٣-٧).

السنة	القيمة	القيمة الإتجاهية (متوسط متحرك لثلاث سنوات)	القيمة الإيجارية (متوسط متحرك لخمس سنوات)
١	ص_1	$\frac{\text{ص}_1 + \text{ص}_2 + \text{ص}_3}{3}$	
٢	ص_2	$\frac{\text{ص}_2 + \text{ص}_3 + \text{ص}_4}{3}$	$\frac{\text{ص}_1 + \text{ص}_2 + \text{ص}_3 + \text{ص}_4 + \text{ص}_5}{5}$
٣	ص_3	$\frac{\text{ص}_3 + \text{ص}_4 + \text{ص}_5}{3}$	$\frac{\text{ص}_2 + \text{ص}_3 + \text{ص}_4 + \text{ص}_5 + \text{ص}_6}{5}$
٤	ص_4	وهكذا	وهكذا
٥	ص_5		
٦	ص_6		

شكل رقم (١٣ – ٧)

كيفية حساب القيم الاتجاهية من المتوسطات المتحركة.

وباستخدام هذا الأسلوب يمكن استنتاج سلسله جديدة تستبدل قيم السلسلة الأصلية بمتوسطات لعدد من الفترات الزمنية. فالمتوسطات الحسابية من الدرجة ٥ (أى على أساس ٥ سنوات) تستبدل قيمة كل سنة للسلسلة الأصلية بمتوسط القيم الخاصة لهذه السنة وقيم السنتين السابقتين لها والسنتين اللاحقتين لها، وبالمثل يمكن تكوين سلسلة زمنية جديدة من القيم هى سلسلة المتوسطات المتحركة لأى عدد من السنوات، وهى قيم اتجاهية أقل عدد من القيم الأصلية، ففى حالة ن = ٣ نجد أنها تبدأ بقيمة أمام السنة الثانية فى السلسله وتنتهى بقيمة أمام السنة

الأخيرة. وعموماً ما كلما كانت المتوسطات المتحركة محسوبة لعدد أكبر من السنوات كلما حصلنا على قيم أكثر تمهيداً.

مثال:

إذا كانت لدينا السلسلة الزمنية التى تمثل الإنتاج السنوى فى دولة ما من الفحم (بمليون الكيلو جرامات) للسنوات من ١٩٤٨ - ١٩٥٨ التى يوضحها الجدول رقم (١٣-٥)، وعلى فرض أن طول الفترة (ن) هو ٥ سنوات فإن المتوسطات المتحركة للكمية المنتجة للفحم للسلسلة المذكورة يمكن حسابها كما فى الجدول التالى:

جدول رقم (١٣-٦)

حساب القيم الاتجاهية بطريقة المتوسطات المتحركة (على أساس ٥ سنوات) لإنتاج السنوى من الفحم (مليون لكيلو جرام) فى دولة ما فى الفترة من ١٩٤٨ - ١٩٥٨

(١) السنة	(٢) السنة	(٣) ٥ سنوات مجموع متحرك	(٤) ٥ سنوات متوسط متحرك
١٩٤٨	٥٠,٠		
١٩٤٩	٣٦,٥		
١٩٥٠	٤٣,٠	٢١٢,٩	٤٢,٦
١٩٥١	٤٤,٥	٢٠١,٠	٤٠,٢
١٩٥٢	٣٨,٩	١٩٧,١	٣٩,٤
١٩٥٣	٣٨,١	١٩٢,٨	٣٩,٦
١٩٥٤	٣٢,٦	١٩٠,٠	٣٨,٠
١٩٥٥	٣٨,٧	١٩٢,٢	٣٨,٤
١٩٥٦	٤١,٧	١٨٧,٩	٣٧,٦
١٩٥٧	٤١,١		
١٩٥٨	٣٣,٨		

بالنظر إلى الجدول السابق نجد أن المجموع المتحرك الأول (٢١٢,٩) بالعمود

الثالث هو المجموع من العنصر الأول إلى العنصر الخامس فى العمود الثانى إلى العنصر السادس وهكذا. ومن الناحية العملية فإنه بعد الحصول على المجموع المتحرك الأول (٢١٢,٩) يمكن الحصول بسهولة على المجموع المتحرك الثانى وذلك بطرح ٥٠,٠ (العنصر الأول فى العمود الثانى) وإضافة ٣٨,١ (العنصر السادس فى العمود الثانى) فتكون النتيجة ٢٠١,٠. وبنفس الطريقة يمكن الحصول على المجاميع المتحركة بسهولة ويسر. وبقسمة كل مجموع متحرك على ٥ ينتج المتوسط المتحرك.

جدول رقم (١٣-٧)

حساب القيم الاتجاهية بطريقة المتوسطات المتحركة (على أساس ٤ سنوات) للإنتاج السنوى من الفحم (مليون كيلو جرام) فى دولة ما فى الفترة ١٩٤٨ - ١٩٥٨

(١) السنة	(٢) البيانات	(٣) ٤ سنوات مجموع متحرك	(٤) ٤ سنوات متوسط متحرك غير مركزى	(٥) ٢ سنة مجموع متحرك للعمود (٤)	(٦) ٤ سنوات م. متحرك مركزى (العمود الخامس ÷٢	(٧) ٢ سنة مجموع متحرك للعمود (٣)	(٨) ٤ سنوات م. متحرك مركزى العمود السابع ÷ ٨
١٩٤٨	٥٠,٠						
١٩٤٩	٣٦,٥						
١٩٥٠	٤٣,٠	١٧٤,٠	٤٣,٥	٨٤,٢	٤٢,١	٣٣٦,٩	٤٢,١
١٩٥١	٤٤,٥	١٦٢,٩	٤٠,٧	٨١,٨	٤٠,٩	٣٢٧,٤	٤٠,٩
١٩٥٢	٣٨,٩	١٦٤,٥	٤١,١	٧٩,٦	٣٩,٨	٣١٨,٦	٣٩,٨
١٩٥٣	٣٨,١	١٥٤,١	٣٨,٥	٧٥,٦	٣٧,٨	٣٠٢,٤	٣٨,٨
١٩٥٤	٣٢,٦	١٤٨,٣	٣٧,١	٧٤,٩	٣٧,٥	٢٩٩,٤	٣٧,٥
١٩٥٥	٣٨,٧	١٥١,١	٣٧,٨	٧٦,٣	٣٨,٢	٣٠٥,٢	٣٨,٢
١٩٥٦	٤١,٧	١٥٤,١	٣٨,٥	٧٧,٣	٣٨,٧	٣٠٩,٤	٣٨,٧
١٩٥٧	٤١,١	١٥٥,٣	٣٨,٨				
١٩٥٨	٣٣,٨						

بالرجوع إلى الجدول السابق نلاحظ أننا حصلنا أولاً على المتوسطات المتحركة غير المركزية (على أساس ٤ سنوات)، أى التى تتركز بين السنوات المتتالية. ولحساب المتوسطات المتحركة لسنوات السلسلة نفسها قمنا بحسب مجموع متحرك للمتوسطات المتحركة (على أساس ٢ سنة) كما فى العمود (٥). ثم حصلنا على المتوسطات المتحركة المركزية عن طريق قسمة بيانات العمود (٥). على ٢ (أى متوسط المتوسطات المتحركة للسنتين المتتاليتين) ويمكن أن نصل إلى نفس النتيجة [بيانات العمود (٦) إذا قمنا بجمع المجموع المتحرك لكل فترتين (العمود (٧)] من العمود (٣) ثم بقسمه هذا المجموع على ٨ (عدد سنوات الفترتين) نحصل المتوسطات المتحركة المركزية [العمود (٨)] لسنوات السلسلة نفسها. وكما نلاحظ أنها نفس القيم الاتجاهية فى العمود (٦).

وكما ذكرنا سالفا أن المشكلة الأساسية للمتوسطات المتحركة تتمثل فى اختيار الفترة المناسبة التى يحسب على أساسها المتوسط. ولما كان الهدف من هذا الأسلوب هو تخليص السلسلة الزمنية من الذبذبات والتغيرات حتى يمكن توضيح الاتجاه العام للسلسلة يجب أن تختا – بعناية– الفترة التى تحقق هذا الهدف.

وبصفة عامة، يؤخذ عن طريقة المتوسطات المتحركة أن البيانات فيها تفقد عند بداية ونهاية السلسلة الزمنية. ففى المثال السابق نبدأ العمل ببيانات مكونة من ١١ رقماً وباستخدام متوسط متحرك من الدرجة ٤ فتنتهى بسبعة أرقام. وتبعاً لذلك إذا حسبت المتوسطات المتحركة على أساس عدد قليل من السنوات فإنها تميل إلى تتبع نفس التغيرات فى بيانات السلسلة الأصلية، أى أنها لاتستطيع إذابة التغيرات أو تمهيد الذبذبات الكبيرة. ومن عيوب المتوسطات المتحركة أيضاً أنها تعطى القيم الاتجاهية فقط دون أن تعطى الصيغة أو المعادلة التى يسير عليها التغير، والتى هى أساس التنبؤ. أو بمعنى آخر أن هذه الطريقة لاتبين معدل النمو الحقيقى للظاهرة قد التحليل. ولكن يحسن استخدام هذه الطريقة كأسلوب لتعيين الاتجاه العام إذا لم يكن الاتجاه على شكل خط مستقيم، وعند ما يكون الغرض هو مجرد دراسة حركة السلسلة الزمنية نفسها دون الاهتمام بالتنبؤ بالقيم الاتجاهية لسنوات ليست موجودة أساساً داخل السلسلة الزمنية قيد الدراسة.

٤- طريقة المربعات الصغرى The Least Squares: تعتبر هذه الطريقة التى سبق دراستها فى الفصل السابق (تحليل الانحدار) أكثر الطرق استخداماً لتوفيق معادلة خط الاتجاه العام للبيانات المشاهدة، إذ أنه بواسطتها يمكن تحديد ثوابت الخط المستقيم (الميل م، والجزء المقطوع جـ) من واقع بيانات السلسلة الزمنية موضع التحليل، وفى هذه الحالة لن تختلف القيم الاتجاهية المحسوبة من خط إلى آخر نظراً لعدم اختلاف ثوابت الخط المستقيم. كما أن من خصائص هذه الطريقة أنها تجعل مجموع مربعات الفروق بين القيم المشاهدة والقيم الاتجاهية المحسوبة من معادلة الخط المستقيم أقل مايمكنت (راجع الفصل السابق تحليل الانحدار).

ويبدأ تعيين الاتجاه العام للسلسلة الزمنية بهذه الطريقة بتمثيل بيانات السلسلة بيانياً وتحديد اتجاهها العام بصفة تقريبية. فإذا أظهر التمثيل البيانى وجود اتجاه عام يأخذ شكل الخط المستقيم كانت معادلته من الدرجة الأولى وصورتها العامة:

ص = م س + جـ .... (١٣-١)

أما إذا كشف التمثيل البيانى عن وجود اتجاه عام يأخذ شكل منحنى (قطع مكافئ مثلاً) كانت معادلته من الدرجة الثانية وصورتها العامة هى:

ص = أ + ب س + جـ س^{٢} .... (١٣-٢)

حيث أ، ب، جـ هى مقادير ثابتة يمكن تقديرها.

وسنهتم هنا بتعيين الاتجاه العام على شكل خط مستقيم فقط لما له من أهمية فى دراسة حركة السلسلة الزمنية فى الماضى والتنبؤ الدقيق فيها بتطور الظاهرة قيد التحليل.

ولحساب معاملات أو ثوابت المعادلة (١٣-١) م ، جـ من البيانات المشاهدة فى السلسلة الزمنية، نستخدم كل من المعادلتين الأساسيتين:

مجـ ص = م مجـ س + ن جـ .... (١٣-٣)

مجـ س ص = م مجـ س^{٢} + جـ مجـ س .... (١٣-٤)

حيث ن تمثل عدد سنوات للسلسلة الزمنية، ص هى قيمة الظاهرة، س هى

الزمن وتسمى طريقة الحساب هذه «بالطريقة المطولة» وذلك لأن نقطة الأصل بالنسبة للزمن تكون عند السنة الأولى فى السلسلة الزمنية.

مثال:

إذا افترضنا أن الإنتاج السنوى من الفحم (مليون كيلو جرام) فى دولة ما للفترة ١٩٤٨ - ١٩٥٨ تمثل فى السلسلة الزمنية فى الجدول رقم (١٣-٥) ولتقدير أحسن القيم للثوابت م، جـ نستخدم بيانات الجدول التالى للتعويض بها فى المعادلتين السابقتين (١٣-٣، ١٣-٤).

جدول رقم (١٣-٨)

حساب خط اتجاه مستقيم بطريقة المربعات الصغرى «الطريقة المطولة»

السنة	ص	س	س ص	س٢
١٩٤٨	٥٠,٠	صفر	صفر	صفر
١٩٤٩	٣٦,٥	١	٣٦,٥	١
١٩٥٠	٤٣,٠	٢	٨٦,٠	٤
١٩٥١	٤٤,٥	٣	١٣٣,٥	٩
١٩٥٢	٣٨,٩	٤	١٥٥,٦	١٦
١٩٥٣	٣٨,١	٥	١٩٠,٥	٢٥
١٩٥٤	٣٢,٦	٦	١٩٥,٦	٣٦
١٩٥٥	٣٨,٧	٧	٢٧١,٩	٤٩
١٩٥٦	٤١,٧	٨	٣٣٣,٦	٦٤
١٩٥٧	٤١,١	٩	٣٦٩,٩	٨١
١٩٥٨	٣٣,٨	١٠	٣٣٨,٠	١٠٠
المجموع	٤٣٨,٩	٥٥	٢١١٠,١	٣٨٥

بالتعويض فى المعادلة (١٣ - ٣) ينتج أن:

٤٣٨,٩ = ٥٥ م + ١١ جـ (حيث ن = ١١) .... (١٣-٥)

وبالتعويض فى المعادلة (١٣-٤) ينتج أن:

٢١١٠,١ = ٣٨٥م + ٥٥جـ .... (١٣-٦)

وبضرب المعادلة (١٣-٥) فى ٥ ينتج أن:

٢١٩٤,٥ = ٢٧٥م + ٥٥جـ .... (١٣-٧)

وبطرح المعادلة (١٣-٧) من المعادلة (١٣-٦) ينتج أن:

$$- ٨٤,٤ = ١١٠م \qquad \therefore م = \frac{-٨٤,٤}{١١٠} = -٠,٧٦٧$$

وبالتعويض عن قيمة م فى المعادلة (١٣-٥) نحصل على قيمة جـ

٤٣٨,٩ = ٥٥ × ٠,٧٦٧ + ١١ جـ

٤٣٨,٩ = ٤٢,١٨٥ + ١١ جـ

$$١١ جـ = ٣٩٦,٧١٥ \qquad \therefore جـ = \frac{٣٩٦,٧١٥}{١١} = ٣٦,٠٦٥$$

وبذلك تكون معاملات أو ثوابت المعادلة (١٣-٥) هى:

م = - ٠,٧٦٧ ، جـ = ٣٦,٠٦٥

وبالتعويض عن م، جـ فى معادلة الخط المستقيم (معادلة رقم (١٣-١)، على أساس نقطة أصل سنة ١٩٤٨، ينتج أن:

ص = ٠٧٦٧، س + ٣٦,٠٦٥

ويمكن تسهيل الحسابات بنقل نقطة الأصل بالنسبة للزمن أو وحدة القياس (س) عند منتصف السلسلة تماماً بحيث يكون عدد قيم الظاهرة التى تسبقها مسار لعدد القيم التى تليها وأخذ انحرافات السنين عن نقطة المنتصف، وبذلك يكون مجـ س = صفر. وتسمى هذه الطريقة «بالطريقة المختصرة» التى تبحث أيضاً فى كيفية إيجاد قيمة كل من م ،جـ للحصول على معادلة الخط المستقيم، كمايلى:

$$\text{م} = \frac{\text{مجـ س ص}}{\text{مجـ س}^{2}} \quad \text{عندما مجـ س = صفر} \;............\; (١٣-٨)$$

$$\text{حـ} = \frac{\text{مجـ ص}}{\text{مجـ ن}} \quad \text{عندما مجـ س = صفر} \;...........\; (١٣-٩)$$

وفى المثال السابق نجد أن السلسلة الزمنية فردية العدد حيث ن = ١١ سنة، وفى هذه الحالة نختار منتصف هذه السلسلة وهى سنة ١٩٥٣ كنقطة أصل. وإذ أخذنا انحرافات السنين الأخرى فى السلسلة عن هذه السنة نحصل على بيانات سهلة مبسطة للزمن (س) كما يتضح من الجدول التالى:

جدول رقم (١٣-٩)

حساب خط اتجاه مستقيم بطريقة المربعات الصغرى

(الطريقة المختصرة)

السنة	ص	س	س ص	س ٢	القيم الإتجاهية $\hat{ش}$
١٩٤٨	٥٠,٠	٥ -	٢٥٠,٠ -	٢٥	٤٣,٧
١٩٤٩	٣٦,٥	٤ -	١٤٦,٠ -	١٦	٤٣,٠
١٩٥٠	٤٣,٠	٣ -	١٢٩,٠ -	٩	٤٢,٢
١٩٥١	٤٤,٥	٢ -	٨٩,٠ -	٤	٤١,٤
١٩٥٢	٣٨,٩	١ -	٣٨,٩ -	١	٤٠,٧
١٩٥٣	٣٨,١	صفر	صفر	صفر	٣٩,٩
١٩٥٤	٣٢,٦	١	٣٢,٦	١	٣٩,١
١٩٥٥	٣٨,٧	٢	٧٧,٤	٤	٣٨,٤
١٩٥٦	٤١,٧	٣	١٢٥,١	٩	٣٧,٦
١٩٥٧	٤١,١	٤	١٦٤,٤	١٦	٣٦,٨
١٩٥٨	٣٣,٨	٥	١٦٩,٠	٢٥	٣٦,١
المجموع	٤٣٨,٩	صفر	-٨٤,٤٠	١١٠	

وبالتعويض فى المعادلتين السابقتين (١٣-٨، ١٣-٩) لاستخراج قيمة كل من م، جـ نجد أن:

$$\text{م} = \frac{-٨٤,٤}{١١٠} = -٠,٧٦٧$$

$$\text{جـ} = \frac{٤٣٨,٩}{١١} = ٣٩,٩$$

وبالتعويض عن م ، جـ فى المعادلة الخطية نجد أن:

معادلة خط الاتجاه العام هى: $\hat{\text{ص}}$ = ٠,٧٦٧٠س + ٣٩,٩ ... (١٣-١٠)

وهى معادلة الخط المستقيم لقيم ص بدلاً من قيم من المختصرة حيث نقطة الأصل سنة ١٩٥٣ ، س مقاسه بوحدات سنوية (١ سنة)، ص هى كمية الإنتاج السنوى.

ونلاحظ أن قيمة م = ٠,٧٦٧٠ التى حسبناها بالطريقة المختصرة هى نفس القيمة التى حصلنا عليها بالطريقة المطلوبة، بينما قيمة جـ (الجزء المقطوع من الاحداثى الصادى) غير متساوية فى الحالتين، والسبب فى ذلك يرجع إلى أن عملية نقل نقطة الأصل لا تؤثر على ميل الخط المستقيم بقدر أنها تنقل نقطة تقاطع خط الاتجاه العام مع المحور الرأسى فى الرسم، إذ تعد نقطة التقاطع فى الحالة الأخيرة هى نقطة الأصل الجديدة، كذلك نلاحظ أن قيمة جـ= ٣٩,٩ هى القيمة الاتجاهية للسنة التى فى منتصف السلسلة وهى سنة ١٩٥٣ حيث س المختصرة لهذه السنة تساوى صفر وهى نقطة الأصل المتغير الزمن س.

ويمكن رسم خط الاتجاه العام بعد ذلك بمعلومية الميل (م) والجزء المقطوع من الاحداثى الصادى (جـ). ولحقيقة أن كل نقطة تقع على هذا الخط إنما تحقق معادلته، فإنه يمكن معرفة أية قيمة من قيم الظاهرة إذا عرفنا قيمة (س) أى الوحدة الزمنية المناظرة. وكل قيمة من قيم الظاهرة التى تقع على خط الاتجاه العام وتناظر وحدة زمنية معينة يطلق عليها اسم «القيمة الاتجاهية» ويرمز لها بالرمز $\hat{\text{ص}}$

والقيم الاتجاهية التى فى العمود الأخير فى الجدول رقم (١٣-٩) حصلنا عليها بالتعويض فى المعادلة السابقة حيث أن القيمة الاتجاهية (صٌ) لسنة ١٩٥٢ مثلاً.

صٌ = -٠,٧٦٧ × (-١) + ٣٩,٩

= ٠,٧٦٧ × ٣٩,٩ = ٤٠,٦٦٧ = ٤٠,٧ تقريباً

وتبعاً للنتائج التى حصلنا عليها يمكن استخدام معادلة خط الاتجاه العام فى التنبؤ بالقيم الاتجاهية لسنوات غير موجودة داخل السلسلة الزمنية نفسها: فمثلاً لحساب القيمة الاتجاهية لإنتاج الفحم فى سنة ١٩٥٩ ، نعوض فى معادلة خط الاتجاه العام بوضع س = ٦

∴ صٌ لسنة ١٩٥٩ = - ٠,٧٦٧ × (٦) + ٣٩,٩

= - ٤,٦٠٢ × ٣٩,٩ = ٠٣٥,٢٩٨

أما إذا كانت السلسلة الزمنية تحتوى على عدد زوجى من السنين فلا يوجد عندئذ سنة فى منتصف السلسلة تجعل مجموع انحرافات بقية السنين عنها تساوى صفراً. ويمكن التغلب على هذه الصعوبة بأخذ منتصف السلسلة بين السنين فى منتصف السلسلة نقل نقطة الأصل إلى منتصف السلسلة. فإذا كانت السلسلة تقع بين الفترة ١٩٤٨ - ١٩٥٧ فإن منتصف سنة ١٩٥٢ يعتبر نقطة الأصل الجديدة والتى تجعل مجـ س = صفر إذا جعلنا وحدة الزمن فى هذه الحالة تساوى نصف السنة.

فإذا كانت لدينا سلسلة من ١٠ سنوات فإنه يمكننا توفيق أحسن خط مستقيم للاتجاه العام كالآتى:

جدول رقم (١٣-١٠)

حساب خط اتجاه مستقيم للسلسلة الزوجية بالطريقة المختصرة

السنة	ص	س	س ص	س ٢
١٩٤٨	٥٠,٠	٩-	٤٥٠,٥ -	٨١
١٩٤٩	٣٦,٥	٧-	٢٥٥,٥ -	٤٩
١٩٥٠	٤٣,٠	٥-	٢١٥,٠ -	٢٥
١٩٥١	٤٤,٥	٣-	١٣٣,٥ -	٩
١٩٥٢	٣٨,٩	١-	٣٨,٩ -	١
١٩٥٣	٣٨,١	١	٣٨,١	١
١٩٥٤	٣٢,٦	٣	٩٧,٨	٩
١٩٥٥	٣٨,٧	٥	١٩٣,٥	٢٥
١٩٥٦	٤١,٧	٨	٢٩١,٩	٤٩
١٩٥٧	٤١,١	٩	٣٦٩,٩	٨١
المجموع	٤٠٥,١	صفر	١٠١,٧-	٣٣٠

من الجدول يتبين أن:

مجـ ص = ٤٠٥,١ ، مجـ س ص = ١٠١,٧ ، مجـ س^{٢} = ٣٢٠

$$\text{م} = \frac{\text{مجـ س ص}}{\text{مجـ س}^{٢}} \quad \text{عندما مجـ س = صفر}$$

$$\therefore \text{م} = \frac{-١٠١,٧}{٣٣٠} = -٠,٣٠٨$$

$$\therefore \text{جـ} = \frac{\text{مجـ ص}}{\text{ن}} \quad \text{عندما مجـ س = صفر}$$

$$\therefore \text{حـ} = \frac{٤٠٥,١}{١٠} = ٤٠,٥١$$

وبالتعويض عن م، جـ في معادلة الخط المستقيم ينتج أن:

$$\hat{ص} = -٠,٣٠٨ \text{ س } + ٤٠,٥١$$

على أساس أن المتوسط الفرضي هو منتصف سنة ١٩٥٢ وأن الوحدة الزمنية هنا هي نصف السنة.

ولتخليص الظاهرة قيد البحث من أثر الاتجاه العام (مع بقاء وقوعها تحت تأثير العوامل الأخرى وهي أثر الدورة، أثر الموسم، أثر التغيرات العشوائية) نقوم بقسمة القيمة الحقيقية للظاهرة على القيمة الاتجاهية لها (أثر الاتجاه العام). وضربها في ١٠٠ حتى نحصل على أثر الاتجاه العام في شكل نسب مئوية، أي أن:

قيمة الظاهرة بعد تخليصها من أثر الاتجاه العام =

$$\frac{\text{القيمة الحقيقية}}{\text{القيمة الاتجاهية}} \times ١٠٠ = \frac{ص}{\hat{ص}} \times ١٠٠$$

وبالرجوع إلى السلسلة الزمنية الموجودة بالجدول رقم (١٣-٩) فإنه يمكننا تخليص البيانات حصلنا عليها باستخدام المعادلة $\hat{ص}$ = ٧٦٧٠ س + ٣٩,٩

بالطريقة المختصرة لطريقة المربعات الصغرى كما هي الحال في الجدول التالي:

ثانياً: تقدير التغيرات الدورية:

تظهر التغيرات الدورية في أوقات الدورات الاقتصادية، أي يقع تأثيرها في فترات النمو أو الانكماش الاقتصادي، وتسبب بذلك ذبذبات في النشاط الاقتصادي. وكما أسفلنا أن الغرض الزساسي من تحليل السلسلة الزمنية هو دراسة طبيعية وأسباب الذبذبات بها وعزل التغيرات - ومنها التغيرات الدورية - التي تؤثر فيها عن طريق قياس أثر هذه التغيرات والتنبؤ بوقوعها، لتفادي التأثيرات الخطرة لها ووضع الحلول للصعوبات التي قد تنجم عنها. وتحدد دورة التغيرات الدورية بالفترة بين قمتين أو بين قاعي موجتين متتاليتين على الرسم البياني للسلسلة الزمنية. وبالرجوع إلى بيانات الجدول رقم (١٣-١١) والشكل رقم (١٣-٨) يمكن

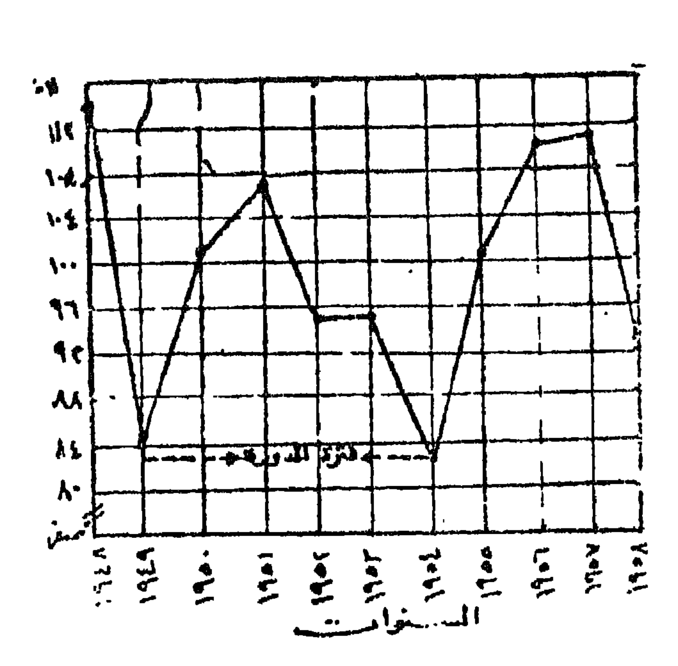
تحديد فترة الدورة بالمسافة الأفقية بين القاعين الممثلين للسنتين ١٩٤٩ ، ١٩٥٤ وهى تمثل دورة طولها خمس سنوات.

جدول رقم (١٣-١١)

تخليص بيانات إنتاج الفحم (مليون كيلو جرام) من أثر الاتجاه العام

السنة	ص	س´	س ٢
١٩٤٨	٥٠,٠	٤٣,٧	١١٤,٤٢
١٩٤٩	٣٦,٥	٤٣,٠	٨٤,٨٨
١٩٥٠	٤٣,٠	٤٢,٢	١٠١,٨٩
١٩٥١	٤٤,٥	٤١,٤	١٠٧,٤٩
١٩٥٢	٣٨,٩	٤٠,٧	٩٥,٥٨
١٩٥٣	٣٨,١	٣٩,٩	٩٥,٤٩
١٩٥٤	٣٢,٦	٣٩,١	٨٣,٣٧
١٩٥٥	٣٨,٧	٣٨,٤	١٠٠,٧٨
١٩٥٦	٤١,٧	٣٧,٦	١١٠,٩٠
١٩٥٧	٤١,١	٣٦,٨	١١١,٦٨
١٩٥٨	٣٣,٨	٣٦,١	٩٣,٦٣

ويكتشف عملية تقرير التغيرات الدورية صعوبة رئيسية وهى أن التأثيرات الدورية لاتعيد نفسها إذ أنها تقع كل فترة طويلة من الزمن وبطريقة غير منتظمة عكس التغيرات الموسمية. ولكن يمكن القول أن دراسة الدورات فى سلسلة زمنية معينة تعبر عن نموذج عام تتشابه فيه الدورات جميعها مع وجود بعض الفروق بينها فى طول الفترة والسعة. وبناء على ذلك فإنه عند تحليل السلسلة الزمنية لتقدير التأثيرات الدورية يجب أن تحدد ما إذا كانت السلسلة الزمنية سنوية أم موسمية.



شكل رقم (١٣-٨) تحديد فترة الدورة

ففى حالة إذا كانت السلسلة الزمنية سنوية فإن قيمة الظاهرة - بعد إهمال تأثير التغيرات العشوائية - تكون تحت تأثير عاملين فقط هما: أثر الاتجاه العام وأثر التغيرات الدورية، لأن العامل الثالث وهو أثر التغيرات الموسمية لايوجد له تأثير إذا كانت السلسلة سنوية وعلى ذلك فإنه يمكن تقدير أثر التغيرات الدورية التى يمكن اعتبارها انحرافات عن القيم الاتجاهية السنوية عن طريق قسمة القيم الأصلية الظاهرة على القيم الاتجاهية لها، إذ أن:

قيمة الظاهرة (ص) = أثر الاتجاه العام (ص) × أثر التغيرات الدورية (د) [مع إهمال أثر التغيرات العشوائية].

فتكون التغيرات الدورية (د) = $\frac{\text{قيمة الظاهرة (ص)}}{\text{القيمة الاتجاهية }(\hat{\text{ص}})}$ ويفضل ضرب هذه

النسبة فى ١٠٠ حتى نحصل على النسبة المئوية للتغيرات الدورية فى السلسلة الزمنية موضع التحليل.

وباستخدام بيانات الجدول رقم (١٣-١٢) تمكنا من حساب أحسن خط مستقيم يمثل الاتجاه العام للسلسلة الزمنية التى تبين إنتاج محصول القمح (ألف طن) فى دولة ما فى الفترة من ١٩٤٦ - ١٩٥٨ بطريقة المربعات الصغرى، والتى أعطتنا المعادلة الآتية (على أساس نقطة الأصل هى سنة ١٩٥٢ والوحدة الزمنية هى سنة):

صˆ = ١,٨٢ س + ٩٢,٤

وباستخدام هذه المعادلة حصلنا على القيم الاتجاهية (صˆ) لسنوات السلسلة عن طريق التعويض بقيم (س) المناظرة وهى الفترة ١٩٤٦ - ١٩٥٨ . وبقيمة القيم الأصلية للظاهرة على القيم الاتجاهية حصلنا على قيمة الظاهرة بعد تخليص البيانات الأصلية من أثر الاتجاه العام، وهذا يعنى بالتالى أثر التغيرات الدورية التى وضعت فى صورة نسب مئوية، حتى تسهل معها عملية المقارنة، ووضعت النتائج فى الجدول التالى:

أما إذا كانت السلسلة الزمنية موسمية فإنه بإهمال أثر التغيرات فإن قيمة لاظاهرة (ص) تكون حينئذ:

ص = صˆ × و × د

وتكون التغيرات الدورية (د) فى هذه الحالة هى:

د = ص / (صˆ × و)

ويعنى ذلك أننا أردنا تقدير أثر التغيرات الدورية فإننا نقوم بقسمة القيم الأصلية للظاهرة على حاصل ضرب القيم الاتجاهية لها فى النسب الموسمية مع إهمال التغيرات العشوائية وبذلك نحصل على مقياس للذبذبات الدورية.

جدول رقم (١٣-١٢)

حساب أثر التغيرات الدورية لإنتاج القمح (ألف طن) في دولة ما

في الفترة من ١٩٤٦ - ١٩٥٨

السنة	الإنتاج (ص)	القيم الاتجاهية (صٓ)	النسبة المئوية للتغيرات الدورية	السنة	الإنتاج (ص)	القيم الاتجاهية (صٓ)	النسبة المئوية للتغيرات الدورية
١٩٤٦	٨٢,٣	٨١,٥	١٠١,٠	١٩٥٣	٩٦,١	٩٤,٢	١٠٢,٠
١٩٤٧	٩٥,٦	٨٣,٣	١١٤,٨	١٩٥٤	٨٩,١	٩٦,٠	٩٢,٨
١٩٤٨	٨١,٣	٨٥,١	٩٥,٥	١٩٥٥	١١٢,٢	٩٧,٩	١١٤,٦
١٩٤٩	٦٩,٤	٨٧,٠	٧٨,٨	١٩٥٦	١٠٠,٤	٩٩,٧	١٠٠,٧
١٩٥٠	٩٢,٨	٨٨,٩	١٠٤,٤	١٩٥٧	١٠٦,٩	١٠١,٥	١٠٥,٣
١٩٥١	٨٣,٤	٩٠,٦	٩٢,١	١٩٥٨	١٩٦,٦	١٠٣,٣	٩٣,٥
١٩٥٢	٨٨,٢	٩٢,٤	٩٥,٥				

وجدير بالذكر أن العديد من الباحثين الاحصائيين ينصح في مثل هذه الحالات بأخذ سلسلة زمنية طويلة للظاهرة قيد التحليل، حتى يمكن أن تظهر ملامح التأثيرات الدورية بوضوح، وفي الوقت نفسه لا يجب إهمال التغيرات العشوائية على الرغم من صعوبة تحديدها أو التنبؤ بها.

ثالثاً: التغيرات الموسمية:

ذكرنا أن التغيرات الموسمية هي إحدى العوامل التي يظهر أثرها على المنحنى التاريخي للسلسلة الزمنية، وأن سبب هذه التغيرات يرجع إلى عوامل طبيعية مثل الظروف المناخية أو الحركات التكتونية السريعة (الزلازل والبراكين)، وعوامل إجتماعية كالعادات والتقاليد والمواسم والأعياد التي تتكرر كل عام. وتفيد دراسة التغيرات الموسمية في معرفة نموذج التغيرات نفسها وعمل المقارنات بين التغيرات

الموسمية فى السنوات المختلفة، ومعرفة إلى أى مدى تؤثر هذه التغيرات فى قيم الظاهرة.

ولحساب أثر التغيرات الموسمية هناك طرق عديدة الغرض الأساسى منها هو معرفة قيمة هذا الأثر على السلسلة الزمنية بعد استبعاد كل العوامل الأخرى المؤثرة، أو بمعنى آخر تحديد المعامل الموسمى كما لو أن قيم الظاهرة لم تتأثر إلا بالتأثير الموسمى فقط. ويطلق على مجموعة الأرقام التى توضح القيم النسبية الموسمية للظاهرة خلال شهور أو فصول السنة إسم «الدليل الموسمى Seasonal Index» للظاهرة. وسنعرض هنا بإختصار لطريقتين من طرق حساب الدليل الموسمى وهما: طريقة متوسط النسب المئوية الموسمية، وطريقة النسبة المئوية للاتجاه العام (أو نسبة الاتجاه العام).

١ - طريقة متوسط النسب المئوية: تستخدم هذه الطريقة فى تقدير التغيرات الموسمية حيث يعبر فيها عن بيانات كل موسم (سواء كان الموسم يوما أو أسبوعاً، أو شهر أو فصلا) خلال السنة كنسبة مئوية من المتوسط فى السنة. ثم نحصل على متوسط النسبة للمواسم المتقابلة فى مختلف السنوات وذلك باستخدام مقياس المتوسط الحسابى أو الوسيط (من الأفضل حذف القيم المتطرفة أو الشاذة عند استخدام المتوسط الحسابى) وتمثل - فى النهاية - النسب المئوية للمواسم «الدليل الموسمى» المطلوب. فإذا كان متوسط الدليل الموسمى أكبر من أو أقل من ١٠٠ ٪ فيجب تعديله بالضرب فى معامل ملائم.

مثال:

يأخذ سلسلة زمنية توضح الطاقة الكهربائية الشهرية (مليون كيلو وات ساعة) المستهلكة فى إضاءة الشوارع والطرق فى دولة ما فى السنوات ١٩٦١ - ١٩٥٨ وجد أن توزيع هذه الطاقة حسب شهور السنة كالآتى:

جدول رقم (١٣-١٣)
الطاقة الكهربائية (مليون كيلو وات ساعة).
المستهلكة فى دولة ما فى الفترة من ١٩٥١ - ١٩٥٨

الشهر / السنة	١٩٥١	١٩٥٢	١٩٥٣	١٩٥٤	١٩٥٥	١٩٥٦	١٩٥٧	١٩٥٨
يناير	٣١٨	٣٤٢	٣٦٧	٣٩٢	٤٢٠	٤٥٣	٤٨٧	٥٢٩
فبراير	٢٨١	٣٠٩	٣٢٨	٣٤٩	٣٧٨	٤١٢	٤٤٠	٤٧٧
مارس	٢٧٨	٢٩٩	٣٢٠	٣٤٢	٣٧٠	٣٩٨	٤٢٩	٤٦٣
ابريل	٢٥٠	٢٦٨	٢٨٧	٣١١	٣٣٤	٣٦٢	٣٩٣	٤٢٣
مايو	٢٣١	٢٤٩	٢٦٩	٢٩٠	٣١٤	٣٤١	٣٧٠	٣٩٨
يونيو	٢١٦	٢٣٦	٢٥١	٢٧٣	٢٩٦	٣٢٢	٣٤٧	٣٨٠
يوليو	٢٢٣	٢٤٢	٢٥٩	٢٨٢	٣٠٥	٣٣٥	٣٥٧	٣٨٩
أغسطس	٢٤٥	٢٦٢	٢٨٤	٣٠٥	٣٣٠	٣٥٩	٢٨٨	٤١٩
سبتمبر	٢٦٩	٢٨٨	٣٠٩	٣٢٨	٣٥٦	٢٩٢	٤١٥	٤٤٨
اكتوبر	٣٠٢	٣٢١	٣٤٥	٣٦٤	٣٩٦	٤٢٧	٤٥٧	٤٩٣
نوفمبر	٣٢٥	٣٤٢	٣٦٧	٣٨٩	٤٢٢	٤٥٤	٤٩١	٥٢٦
ديسمبر	٣٤٧	٣٦٤	٣٩٤	٤١٧	٤٥٢	٤٨٣	٥١٦	٥٦٠
المجموع	٢٢٨٥	٣٥٢٢	٣٧٨٠	٤٠٤٢	٤٣٧٣	٤٧٣٨	٥٠٩٠	٥٥٠٥
المتوسط	٢٧٣,٧	٢٩٣,٥	٣١٥٠	٣٣٦٠٨	٣٦٤٠٤	٢٩٤,٨	٤٢٤,٢	٤٥٨,٧

ولتقدير الدليل الموسمى بطريقة متوسط النسب المئوية بقسمة البيانات الشهرية المعطاه على المتوسطات الشهرية المقابلة لكل سنة مع التغير عن النتيجة كنسبة مئوية فنحصل على البيانات التى يوضحها الجدول التالى:

جدول رقم (١٣-١٦)

النسب المئوية والدليل الموسمي للطاقة الكهربائية المستهلكة

في دولة ما في الفترة من ١٩٥١ - ١٩٥٨

السنة / الشهر	١٩٥١	١٩٥٢	١٩٥٣	١٩٥٤	١٩٥٥	١٩٥٦	١٩٥٧	١٩٥٨	المجموع	المتوسط
يناير	١١٦,٢	١١٦,٥	١١٦,٥	١١٦,٤	١١٥,٣	١١٤,٧	١١٤,٨	١١٥,٣	٩٢٥,٧	١١٥,٧
فبراير	١٠٢,٧	١٠٥,٣	١٠٤,١	١٠٣,٦	١٠٣,٧	١٠٤,٤	١٨٣,٧	١٠٤,٠	٨٣١,٥	١٠٣,٩
مارس	١٠١,٦	١٠١,٩	١٠١,٦	١٠١,٥	١٠١,٥	١٠٠,٨	١٠١,١	١٠٠,٩	٨١٠,٩	١٠١,٤
ابريل	٩١,٣	٩١,٣	٩١,١	٩٢,٣	٩١,٧	٩١,٧	٩٢,٦	٩٢,٢	٨٢٤,٢	٩١,٨
مايو	٨٤,٤	٨٤,٨	٨٥,٤	٨٦,١	٨٦,٢	٨٦,٤	٨٧,٢	٨٦,٨	٦٨٧,٣	٨٥,٩
يونيو	٧٨,٩	٨٠,٤	٧٩,٧	٨١,١	٨٢,٢	٨١,٦	٦١,٨	٨٢,٨	٦٤٧,٥	٨٠,٩
يوليو	٨١,٥	٨٢,٥	٨٢,٢	٨٣,٧	٨٣,٧	٨٤,٩	٨٤,٢	٨٤,٨	٦٦٧,٥	٨٣,٤
اغسطس	٨٩,٥	٨٩,٣	٩٠,٢	٩٠,٦	٩٠,٦	٩٠,٩	٩١,٥	٩١,٢	٧٢٣,٩	٩٠,٥
سبتمبر	٩٨,٣	٩٨,١	٩٨,١	٩٧,٤	٩٧,٧	٩٩,٣	٩٧,٨	٩٧,٧	٧٨٤,٤	٩٨,١
اكتوبر	١١٠,٣	١٠٩,٤	١٠٩,٥	١٠٨,١	١٠٨,٧	١٠٨,٢	١٠٧,٧	١٠٧,٥	٨٦٩,٤	١٠٨,٧
نوفمبر	١١٨,٧	١١٦,٥	١١٦,٥	١١٥,٥	١١٥,٨	١١٥,٠	١١٥,٧	١١٤,٧	٩٢٨,٤	١١٦,١
ديسمبر	١٢٦,٨	١٢٤,٠	١٢٥,١	١٢٣,٨	١٢٤,٠	١٢٢,٣	١٢١,٦	١٢٢,١	٩٨٩,٧	١٢٣,٧

فعلى سبيل المثال القيمة الأولى يناير ١٩٥١ في الجدول السابق تحسب كمايلي:

$$\frac{٣١٨}{٢٧٣,٦} \times ١٠٠ = ١١٦,٢$$ ..... وهكذا لباقي شهور نفس السنة.

وجدير بالذكر أنه إذا كانت الشهور كلها متساوية من حيث طاقة الاضاءة المستخدمة لانعدمت التغيرات الرسمية، ولو وجدنا أن متوسط كل شهر يساوى المتوسط العام ولكانت نسبة موسمية تساوى ١٠٠ ٪. وحيث أننا وجدنا أن النسبة الموسمية للشهر الأول هى ١١٥,٧ فمعنى هذا أن متوسط الطاقة الكهربائية

المستخدمة فى هذا الشهر يزيد عن المتوسط العام بمقدار ١٥,٧ ٪، بينما نجد أن النسبة الموسمية لشهر يوليو هى ٨٠,٩ ٪ ومعنى هذا أن متوسط الطاقة الكهربائية المستخدمة فى هذا الشهر ينقص عن المتوسط العام بمقدار ١٩,١ ٪، ولو جمعنا الزيادة والنقص فى النسب المئوية وكان الناتج أكبر من الصفر فيجب تعديل النسب المئوية للشهور بضربها فى معامل ملائم. وبما أن مجموع متوسط النسب المئوية لكل الشهور هو ١٢٠٠,١ ٪ وهو قريب جداً من المجموع المطلوب ١٢٠٠ ٪ بمتوسط ١٠٠ ٪ لشهور السنة جميعها، فإنه ليس من الضرورى تعديل النسب المئوية للشهور والتى تعبر فى هذه الحالة عن «الدليل الموسمى» المطلوب.

ويمكن إستخدام طريقة أخرى لحساب متوسط النسب المئوية تقوم على أساس حساب متوسط عام (متوسط المتوسطات) للمواسم ـ سواء كانت على أساس يومى أو أسبوعى أو شهرى ... الخ ـ عن طرق حساب متوسط قيم كل موسم من مواسم السلسلة الزمنية موضع التحليل، ثم بعد ذلك يحسب المتوسط الحسابى لمتوسطات المواسم وهو مايعبر عنه بالمتوسط العام. وبقسمة المتوسط الحسابى لكل موسم على المتوسط العام وضرب الناتج فى ١٠٠ نحصل على نتائج فى شكل نسب مئوية هى ما يطلق عليه اسم «النسب الموسمية». ويمكن توضيح ذلك بالصورة الرمزية الآتية:

	السنة (١)	السنة (٢)	السنة (٣)	السنة (٤)
الموسم (١)	$س_{١١}$	$س_{٢١}$	$س_{٣١}$	$س_{٤١}$
الموسم (٢)	$س_{١٢}$	$س_{٢٢}$	$س_{٣٢}$	$س_{٤٢}$
الموسم (٣)	$س_{١٣}$	$س_{٢٣}$	$س_{٣٣}$	$س_{٤٣}$

ويكون:

$$\text{متوسط الموسم الأول للسنوات الأربع} = \frac{س_{١١} + س_{٢١} + س_{٣١} + س_{٤٤}}{٤} = \bar{س}_{١}$$

$$\text{متوسط الموسم الأول للسنوات الأربع} = \frac{س_{١٢} + س_{٢٢} + س_{٣٢} + س_{٤٢}}{٤} = \bar{س}_{٢}$$

متوسط الموسم الثالث للسنوات الأربع $= \frac{س_{١٣} + س_{٢٣} + س_{٣٣} + س_{٤٣}}{٤} = \bar{س}_{٣}$

والمتوسط العام للمواسم $= \frac{\bar{س}_{١} + \bar{س}_{٢} + \bar{س}_{٣}}{٣} = \bar{م}$

وعليه فإن:

نسبة الموسم الأول $= \frac{\bar{س}_{١}}{\bar{م}} \times ١٠٠$ ، نسبة الموسم الثانى $= \frac{\bar{س}_{٢}}{\bar{م}} \times ١٠٠$

نسبة الموسم الثالث $= \frac{\bar{س}_{٣}}{\bar{م}} \times ١٠٠$

وتمثل النسب المئوية للمواسم «الدليل الموسمى» للسلسلة الزمنية

وإذا طبقت هذه الطريقة على بيانات الجداول رقم (١٣-١٣) فإننا نحصل على نسب مئوية مساوية تقريباً للنسب المئوية فى الجدول رقم (١٣-١٤)، وهو ما يوضحه الجدول التالى:

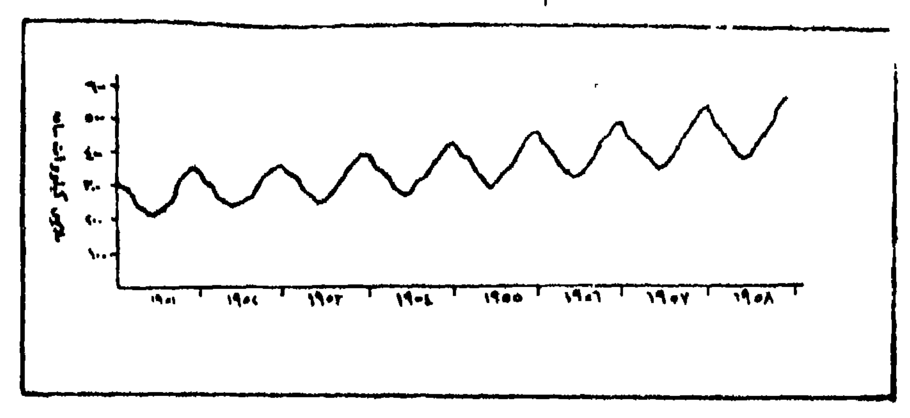
جدول رقم (١٣-١٥)

حساب النسب المئوية الموسمية بطريقة المتوسط العام لبيانات الطاقة الكهربائية بالجدول (١٣-١٣)

الشهر	المجموع	المتوسط	النسب الموسمية	الشهر	المجموع	المتوسط	النسبة الموسمية
يناير	٣٣٠٨	٤١٣,٥	١١٥,٦	أغسطس	٢٥٩٢	٣٢٤,٠	٩٠,٦
فبراير	٢٩٧٤	٣٧١,٧	١٠٣,٩	سبتمبر	٢٨٠٥	٣٥٠,٦	٩٨,٨
مارس	٢٨٩٩	٣٦٢,٤	١٠١,٣	أكتوبر	٣١٠٥	٣٨٨,١	١٠٨,٥
ابريل	٢٦٢٨	٣٢٨,٥	٩١,٩	نوفمبر	٣٣١٦	٤١٤,٥	١١٥,٩
مايو	٢٤٦٢	٣٠٧,٧	٨٦,٠	ديسمبر	٣٥٣٣	٤٤١,٦	١٢٣,٥
يونيو	٢٣٢١	٢٩٠,١	٨١,١	المجموع		٤٢٩١,٧	
يوليو	٢٣٩٢	٢٩٩,٠	٨٣,٦	المتوسط العام		٣٥٧,٦	

٢ – طريقة النسبة المئوية للاتجاه العام: تهدف هذه الطريقة إلى استبعاد التغيرات الدورية والعشوائية من البيانات. وكما لاحظنا أن معظم بيانات الظواهر تتأثر بالاتجاه العام لذلك يجب استبعاد هذا الأثر أيضاً قبل تقدير التأثير الموسمى. ويعبر فى هذه الطريقة عن بيانات الموسم كنسبة مئوية من القيم الاتجاهية فى الموسم. وباستخدام متوسط ملائم لهذه النسب للمواسم المتقابلة نحصل على الدليل الموسمى المطلوب. وكما فى طريقة متوسط النسب المئوية تعدل النسب المئوية التى تحصل عليها إذا لم يكن متوسطها ١٠٠٪.

فمثلاً إذا استخدمنا المتوسطات الشهرية المقابلة لكل سنة من سنوات السلسلة الزمنيـة (٥١ – ١٩٥٨) فى الجـدول رقم (١٣–١٣) للحـصـول على الدليل الموسمى باستخدام طريقة النسبة المئوية للاتجاه العام، يجب استخدام طريقة المربعات الصغرى – الاسبق شرحها – للحصول على القيم الاتجاهية. فمن الرسم البيانى (شكل رقم ١٣–٩) للبيانات الأصلية يتضح أن الاتجاه العام طويل الذى يمكن تقريبه بصورة ماسبة بخط مستقيم.



شكل رقم (١٣–٩)

تطور استخدام الطاقة الكهربائية (مليون كيلو وات ساعة)

فى دولة ما فى الفترة ١٩٥١ – ١٩٥٨

وبافتراض أن القيم الشهرية فى الجدول رقم (١٣–١٣) تقابل منتصف الشهر فإن المتوسطات السنوية فى هذا الجدول تقابل ٣٠ يونيو أو ١ يوليو للسنة القابلة

لكل متوسط، وبأخذ وحدة الزمن فى هذه الحالة نصف سنة، ونقطة الأصل هى ٣١ ديسمبر ١٩٥٤، أو يناير ١٩٥٥، فإنه يمكن توفيق أحسن خط مستقيم للاتجاه العام كالآتى:

جدول رقم (١٣-١٦)
حساب خط الاتجاه العام بطريقة المربعات الصغرى (الطريقة المختصرة)

السنة	ص	س	س ص	$س^2$
١٩٥١	٢٧٣,٧	-٧	-١٩١٥,٩	٤٩
١٩٥٢	٢٩٣,٥	-٥	-١٤٦٧,٥	٢٥
١٩٥٣	٣١٥,٠	-٣	-٩٤٥,٠	٩
١٩٥٤	٣٣٦,٨	-١	-٣٣٦,٨	١
١٩٥٥	٣٦٤,٤	١	٣٣٦٤,٤	١
١٩٥٦	٣٩٤,٨	١	١١٨٤,٤	٩
١٩٥٧	٤٢٤,٢	٥	٢١٢١,٠	٢٥
١٩٥٨	٤٥٨,٧	٧	٣٢١٠,٩	٤٩
المجموع	٢٨٦١,١	صفر	٢٢١٥,٥	١٦٨

وحيث أن معادلة الخط المستقيم هى = ص = م س + حـ،

وبالتعويض عن م بالقيمة $\left(\frac{\text{مجـ س ص}}{\text{مجـ س}^2}\right)$، حـ بالقيمة $\left(\frac{\text{مجـ ص}}{\text{ن}}\right)$ ينتج أن:

$$\text{ص}' = \left(\frac{\text{مجـ س ص}}{\text{مجـ س}^2}\right)\text{س} + \left(\frac{\text{مجـ ص}}{\text{ن}}\right) = \left(\frac{٢٢١٥,٥}{١٦٨٠}\right)\text{س} + \left(\frac{\text{مجـ ص}}{\text{ن}}\right)$$

= ١٣,١٨٨ س + ٣٥٧,٦.

ومن هذه المعادلة نستنتج أن قيم (ص) تزيد ١٣,١٨٨ كل نصف سنة أو ٢,٢ كل شهر(١٣,١٨٨ ÷ ٦)، وبالتالى فعند س = صفر (١ يناير ١٩٥٥) فإن

ص = ٣٥٧,٦، وبعد نصف شهر (١٥ يناير ١٩٥٥) فإن قيمة (ص) تصبح ٣٥٨,٧ (٢,٢ × ١/٢ + ٣٥٧,٦) وهى القيمة الاتجاهية (ص) المقابلة لشهر يناير ١٩٥٥ . وبإضافة المتتالية ٢,٢ إلى ٣٥٨,٧ نحصل على القيمة الاتجاهية لشهر فبراير ١٩٥٥ وهى ٣٦٠,٩ (٣٥٨,٧ + ٢,٢). ولشهر مارس ١٩٥٥ تكون القيمة الاتجاهية (ص) هى ٣٦٣,١ (٣٦٠,٩ + ٢,٢)، وهكذا. وبنفس الطريقة فإنه بالطرح المتتالى للقيمة ٢,٢ من القيمة ٣٥٨,٧ فإننا نحصل على القيمة الاتجاهية لشهر ديسمبر ١٩٥٤ ونوفمبر ١٩٥٤ وهى: ٣٥٦,٥ (٣٥٨,٧ – ٢,٢) و ٣٥٤,٣ (٢٥٦,٥ – ٢,٢) على الترتيب. وبصورة مشابهة نحصل على القيم الاتجاهية الشهرية الموضحة بالجدول التالى:

جدول رقم (١٣-١٧)
القيم الاتجاهية لتطور الطاقة الكهربائية
المستهلكة فى دولة ما فى الفترة من ١٩٥١ – ١٩٥٨

الشهر / السنة	١٩٥١	١٩٥٢	١٩٥٣	١٩٥٤	١٩٥٥	١٩٥٦	١٩٥٧	١٩٥٨
يناير	٢٥٣,١	٢٧٩,٥	٣٠٥,٩	٣٣٢,٣	٣٥٨,٧	٣٨٥,١	٤١١,٥	٤٣٧,٩
فبراير	٢٥٥,٣	٢٨١,٧	٣٠٨,١	٣٣٤,٧	٣٦٠,٩	٣٨٧,٣	٤١٣,٧	٤٤٠,١
مارس	٢٥٧,٥	٢٨٣,٩	٣١٠,٣	٣٣٦,٧	٣٦٣,١	٣٨٩,٥	٤١٥,٩	٤٤٢,٣
ابريل	٢٥٩,٧	٢٨٦,١	٣١٢,٥	٣٣٨,٩	٣٦٥,٣	٣٩١,٧	٤١٨,١	٤٤٤,٥
مايو	٢٦١,٩	٢٨٨,٣	٣١٤,٧	٣٤١,١	٣٦٧,٥	٣٩٣,٩	٤٢٠,٣	٤٤٦,٧
يونيو	٢٦٤,١	٣٩,٥	٣١٦,٩	٣٤٣,٣	٣٦٩,٧	٣٩٦,١	٤٢٢,٥	٤٤٨,٩
يوليو	٢٦٦,٣	٢٩٢,٧	٣١٩,١	٣٤٥,٥	٣٧١,٩	٣٩٨,٣	٤٢٤,٧	٤٥١,١
أغسطس	٢٦٨,٥	٢٩٤,٩	٣٢١,٣	٣٤٧,٧	٣٧٤,١	٤٠٠,٥	٤٢٦,٩	٤٥٣,٣
سبتمبر	٢٧٠,٧	٢٩٧,١	٣٢٣,٥	٣٤٩,٩	٣٧٦,٣	٤٠٢,٧	٤٢٩,١	٤٥٥,٥
اكتوبر	٢٧٢,٩	٢٩٩,٣	٣٢٥,٧	٣٥٢,١	٣٧٨,٥	٤٠٤,٩	٤٣١,٣	٤٥٧,٥
نوفمبر	٢٧٥,١	٣٠١,٥	٣٢٧,٩	٣٥٤,٣	٣٨٠,٧	٤٠٧,١	٤٣٣,٥	٤٥٩,٩
ديسمبر	٢٧٧,٣	٣٠٣,٧	٣٣٠,١	٣٥٦,٥	٣٨٢,٩	٤٠٩,٣	٤٣٥,٧	٤٦٢,١

وبقسمة كل قيمة من القيم الشهرية الأصلية بالجدول رقم (١٣-١٣) على القيمة الاتجاهية بالجدول السابق (جدول رقم ١٣ - ١٧) وبضربها فى ١٠٠ نحصل على نتائج فى شكل نسب مئوية. فعلى سبيل المثال تحسب القيمة الأولى بالجدول كالآتى $\frac{٣١٨}{٢٥٣,١} \times ١٠٠ = ١٢٥,٦$ ٪، وبالمثل يمكن حساب بقيمة القيم فنحصل على القيم الموضحة بالجدول التالى:

جدول رقم (١٣-١٨)

النسب الموسمية للطاقة الكهربائية المستهلكة فى دولة ما فى الفترة من ١٩٥١ - ١٩٥٨

الشهر / السنة	١٩٥١	١٩٥٢	١٩٥٣	١٩٥٤	١٩٥٥	١٩٥٦	١٩٥٧	١٩٥٨
يناير	١٢٥,٦	١٢٢,٤	١٢٠,٠	١١٨,٠	١١٧,١	١١٧,٦	١١٨,٣	١٢٠,٨
فبراير	١١٠,١	١١٠,٠	١٠٦,٥	١٠٤,٣	١٠٤,٧	١٠٦,٤	١٠٦,٤	١٠٨,٤
مارس	١٠٨,٠	١٠٥,٣	١٠٣,١	١٠١,٦	١٠١,٩	١٠٢,٢	١٠٣,١	١٠٤,٧
ابريل	٩٦,٣	٩٣,٧	٩١,٨	٩١,٨	٩١,٤	٩٢,٤	٩٤,٠	٩٥,٢
مايو	٨٨,٢	٨٦,٤	٨٥,٥	٨٥,٠	٨٥,٢	٨٦,٦	٨٨,٠	٨٩,١
يونيو	٨١,٨	٨١,٢	٨٩,٢	٧٩,٥	٨٠,١	٨١,٣	٨٢,١	٨٤,٧
يوليو	٨٣,٧	٨٢,٧	٨١,٢	٨١,٦	٨٢,٠	٨٤,١	٨٤,١	٨٦,٢
أغسطس	٩١,٢	٨٨,٨	٨٨,٤	٨٧,٧	٨٨,٢	٨٩,٦	٩٠,٩	٩٢,٤
سبتمبر	٩٩,٤	٩٦,٩	٩٥,٥	٩٣,٧	٩٤,٦	٩٧,٣	٩٦,٧	٩٨,٤
اكتوبر	١١٠,٧	١٠٧,٣	١٠٥,٩	١٠٣,٤	١٠٤,٦	١٠٥,٥	١٠٦,٠	١٠٧,٧
نوفمبر	١١٨,١	١١٣,٤	١١١,٩	١٠٩,٨	١١٠,٨	١١١,٥	١١٣,٣	١١٤,٤
ديسمبر	١٢٥,١	١١٩,٩	١١٩,٤	١١٧,٠	١١٨,٠	١١٨,٠	١١٨,٤	١٢١,٢

وإذا قورنت المتوسطات الشهرية ببعض فى السلسلة للزمنية قيد التحليل فإن الذبذبات فى هذه المتوسطات الشهرية ترجع إلى التغيرات الموسمية لأن التغيرات

الأخرى تستبعد تلقائياً خصوصاً إذا كان عدد سنين السلسلة كبيراً. ويجب استبعاد القيم المتطرفة فى السلسلة الزمنية والتى تكون بسبب الظواهر الفجائية، ولذلك يفضل استخدام الوسيط بدلاً من المتوسط الحسابى، لأن الوسيط لا يتأثر بالقيم المتطرفة. وفى هذه الحالة ترتب النسب للسنتين فى المنتصف (راجع كيفية حساب الوسيط فى حالة التوزيعات غير المبوبة ذات الأعداد الزوجية من القيم). وأخيراً تنسب وسيط كل شهر إلى الوسيط الشهرى العام (وهو عبارة عن مجموع الوسيطات الشهرية ١١٩٦,١ مقسوماً على ١٢، أى يساوى ٩٩,٧ ٪). ونظراً لأن مجموع قيم الوسيطات الشهرية وهو ١١٩٦,١ أقل من المجموع الذى يجب أن يكون وهو ١٢٠٠ فإننا نعدل هذه القيم بضربها فى النسبة ١١٩٦,١ ÷ ١٢٠٠. وبهذه الطريقة نحصل على الدليل الموسمى المطلوب كما هو موضح بالجدول التالى:

جدول رقم (١٣-١٩)

الوسيط والدليل الموسمى للبيانات السابق عرضها فى الجدول رقم (١٣-١٨)

الشهر	الوسيط	الدليل الموسمى	الشهر	الوسيط	الدليل الموسمى	الشهر	الوسيط	الدليل الموسمى
يناير	١١٩,٢	١١٩,٦	مايو	٨٦,٥	٨٦,٨	سبتمبر	٩٦,٨	٩٧,١
فبراير	١٠٦,٤	١٠٦,٧	يونيو	٨١,٢	٨١,٥	أكتوبر	١٠٦,٠	١٠٦,٣
مارس	١٠٣,١	١٠٣,٤	يوليو	٨٣,٢	٨٣,٥	نوفمبر	١١٢,٦	١١٣,٠
إبريل	٩٣,٠	٩٣,٣	أغسطس	٨٩,٢	٨٩,٥	ديسمبر	١١٨,٩	١١٩,٣

ومما تجدر ملاحظته على الجدول السابق أن أرقام الدليل الموسمى فى الاشهر السبعة الأولى أكبر من نظائرها التى حصلنا عليها بطريقة متوسط النسب المئوية (راجع الجدول رقم (١٣-١٤)، بينما تكون أرقام الدليل الموسمى فى الاشهر الخمسة الأخيرة أقل من مثيلتها فى الجدول رقم (١٣-١٤).

ومن أهم مزايا قياس التغيرات الموسمية هو استخدامها في التنبؤ بمقدار التأثيرات الموسمية في السنوات اللاحقة لسنوات السلسلة الزمنية قيد التحليل. فإذا ما أمكن التنبؤ بالقيمة الاتجاهية في سنة معينة من سنوات السلسلة، فإنه يمكن التنبؤ بقيمة كل موسم في تلك السنة على حدة إذا كانت النسب المئوية محسوبة. ولنفرض أننا نريد أن نتنبأ بمتوسط الطاقة الكهربائية المستخدمة في الاضاءة لشهور سنة ١٩٥٩ باستخدام المعادلة الخطية ص = ١٣,٨٨ س + ٣٥٧,٦ حيث س تمثل وحدة الزمن وهي نصف السنة. وبطريقة إضافة المتتالية ٢,٢ (المعدل الشهرى الذى حصلنا عليه من قسمة المعدل نصف السنوى ١٣,١٨٨ على ٦) إلى القيمة الاتجاهية لشهر ديسمبر ١٩٥٨ وهي ٤٦٢,١ نحصل على القيمة الاتجاهية لشهر يناير ١٩٥٩ وهي ٤٦٤,٣ ٤٦,١٠ + ٢,٢)، وبإضافة المتتالية إلى القيمة الاتجاهية الأخيرة نحصل على القيمة الاتجاهية لشهر فبراير ١٩٥٩ وهي ٤٦٦,٥ (٤٦٤,٣ + ٢,٢)، ولشهر مارس ١٩٥٩ وهي ٤٦٨,٧ (٤٦٦,٥ + ٢,٢) ولشهر إبريل ٤٧٠,٩ ومايو ٤٧٣,١ ويونيو ٤٧٥,٣ ويوليو ٤٧٧,٥ وأغسطس ٤٧٩,٧ وسبتمبر ٤٨١,٩ ، وأكتوبر ٤٨٤,١ ، ونوفمبر ٤٨٦,٣ ، وديسمبر ٤٨٨,٥ . والقيم الاتجاهية السابقة يمكن اعتبارها تقريراً لمتوسط الطاقة الكهربائية لشهور سنة ١٩٥٩ إذا لم يكن هناك تغيرات موسمية. ولكن في دراستنا التغيرات الموسمية استنتجنا أن الدليل الموسمى الذى حصلنا عليه باستخدام طريقة النسب المئوية للاتجاه العام على أساس الوسيط – لشهر يناير مثلاً هو ١١٩,٦٪ لو لم يكن هناك تغيرات موسمية وبافتراض أن الدليل الموسمى سيظل كما هو خلال عام ١٩٥٩ ، فإن التنبؤ بالمتوسط الشهرى للطاقة الكهربائية يجب أن يعدل بأن نأخذ ١١٩,٦٪ من القيمة الاتجاهية المقدرة لشهر يناير ١٩٥٩ ، ١٠٦,٧٪ من القيمة الاتجاهية المقدرة لشهر فبراير ١٩٥٩ .. وهكذا لباقي الشهور كما هو موضح بالجدول التالي:

جدول رقم (١٣-٢٠)

القيم الاتجاهية وأثر الموسمية ومقدار الطاقة الكهربائية المقدرة لسنة ١٩٥٩

الشهر	القيمة الإتجاهية	أثر الموسمية	الطاقة المقدرة ١٩٥٩	الشهر	القيمة الإتجاهية	أثر الموسمية	الطاقة المقدرة ١٩٥٩
يناير	٤٦٤,٣	١,١٩٦	٥٥٥,٣	يوليو	٤٧٧,٥	,٨٣٥	٣٩٨,٧
فبراير	٤٦٦,٥	١,٠٦٧	٤٩٧,٧	أغسطس	٤٧٩,٧	,٨٩٥	٤٢٩,٢
مارس	٤٦٨,٨	١,٠٣٤	٤٨٤,٦	سبتمبر	٤٨١,٩	,٩٧١	٤٦٧,٩
إبريل	٤٧٠,٩	,٩٣٣	٤٣٩,٣	أكتوبر	٤٨٤,١	١,٠١٣	٥١٤,٦
مايو	٤٧٣,١	,٨٦٨	٤١٠,٦	نوفمبر	٤٨٦,٣	١,١٣٠	٥٤٩,٥
يونيو	٤٧٥,٣	,٨١٥	٣٨٧,٤	ديسمبر	٤٨٨,٥	١,٩٨	٥٨٢,٨

ويمكن أخيراً تخليص بيانات الظاهرة من أثر التغيرات الموسمية بنفس طريقة تخليص بيانات الظاهرة من أثر الاتجاه العام وذلك بقسمة كل عنصر في البيانات الأصلية (جدول رقم ١٣-١٣) على أرقام الدليل الموسمي (أو الأرقام القياسية الموسمية) المقابلة. وتسمى البيانات التي نحصل عليها بيانات لا موسمية أو بيانات معدلة لاستبعاد التغيرات الموسمية. مثل هذه البيانات تتضمن أثر الاتجاه العام والتغيرات الدورية والتغيرات العشوائية، أي أن:

$$\frac{\text{القيمة الأصلية (ص)}}{\text{التغيرات الموسمية (و)}} = \text{القيمة الاتجاهية (ص)} \times \text{التغيرات الدورية (د)}$$

× التغيرات العشوائية (ش)

فمثلاً إذا أردنا معرفة متوسط الطاقة الكهربائية المستهلكة (مليون كيلو وات

ساعة) لشهر يناير ١٩٥١ بعد استبعاد التأثيرات الموسمية فإن هذا المتوسط يكون على النحو التالي:

$$\frac{٣١٨ \times ١٠٠}{١١٩,٦} = ٢٦٥,٩$$ مليون كيلو وات ساعة

وبالمثل لباقى قيم الظاهرة فى شهور السنوات المختلفة للسلسلة الزمنية من ١٩٥١ - ١٩٥٨.

رابعاً: تقدير التغيرات العشوائية:

بعد أن تمكننا من استبعاد أثر الاتجاه العام والتغيرات الدورية والتغيرات الموسمية من السلسلة الزمنية، فإنها سوف تقع تحت تأثير التغيرات العشوائية أو التغيرات الفجائية (مثل وقوع الزلزال والبراكين والحروب) التى من الصعب التكن بها أو التنبؤ بوقوعها، وبالتالى يصعب تحديد حجم هذه التغيرات وتحديد مدى تأثيرها فى قيمة الظاهرة. ويمكن تقدير قوة تأثير التغيرات العشوائية نظرياً حيث أننا سوف لانحتاج إلا إلى مقارنة القيم الزصلية بالقيم النظرية المحسوبة على أساس خط الاتجاه العام والتغيرات الموسمية، فأى فرق أو إنحراف بين القيمة الأصلية والقيمة النظرية تنسبه إلى التغيرات العشوائية. كما يمكن حساب التغيرات العشوائية بغرض استبعادها وحذفها من قيم الظاهرة وذلك بقسمة البيانات الأصلية (ص) للظاهرة على حاصل ضرب كل من أثر الاتجاه العام والتغيرات الدورية فقط فى حالة إذا كانت السلسلة الزمنية موضع التحليل سنوية، وعلى حاصل ضرب كل من أثر الاتجاه العام والتغيرات الدورية والموسمية إذا كانت السلسلة الزمنية موسمية سنوية، أى أن:

(١) فى حالة السلسلة الزمنية السنوية:

التغيرات العشوائية (ش) = $$\frac{ص}{ت \times د}$$

(٢) فى حالة السلسلة الزمنية الموسمية:

التغيرات العشوائية (ش) = $\frac{\text{ص}}{\text{ت} \times \text{د} \times \text{و}}$

ومن الناحية العملية وجد أن التغيرات العشوائية (غير المنتظمة) تتجه إلى أن تكون ذات تأثير قليل (حجم صغير)، وأنها غالباً تتجه إلى أن تتبع نمط التوزيع المعتدل (الطبيعي)، أى أنها عبارة عن انحرافات صغيرة تحدث بتكرارات كبيرة أما الانحرافات الكبيرة فتحدث بتكرارات صغيرة.

المحتويات

فهرس محتويات الكتاب

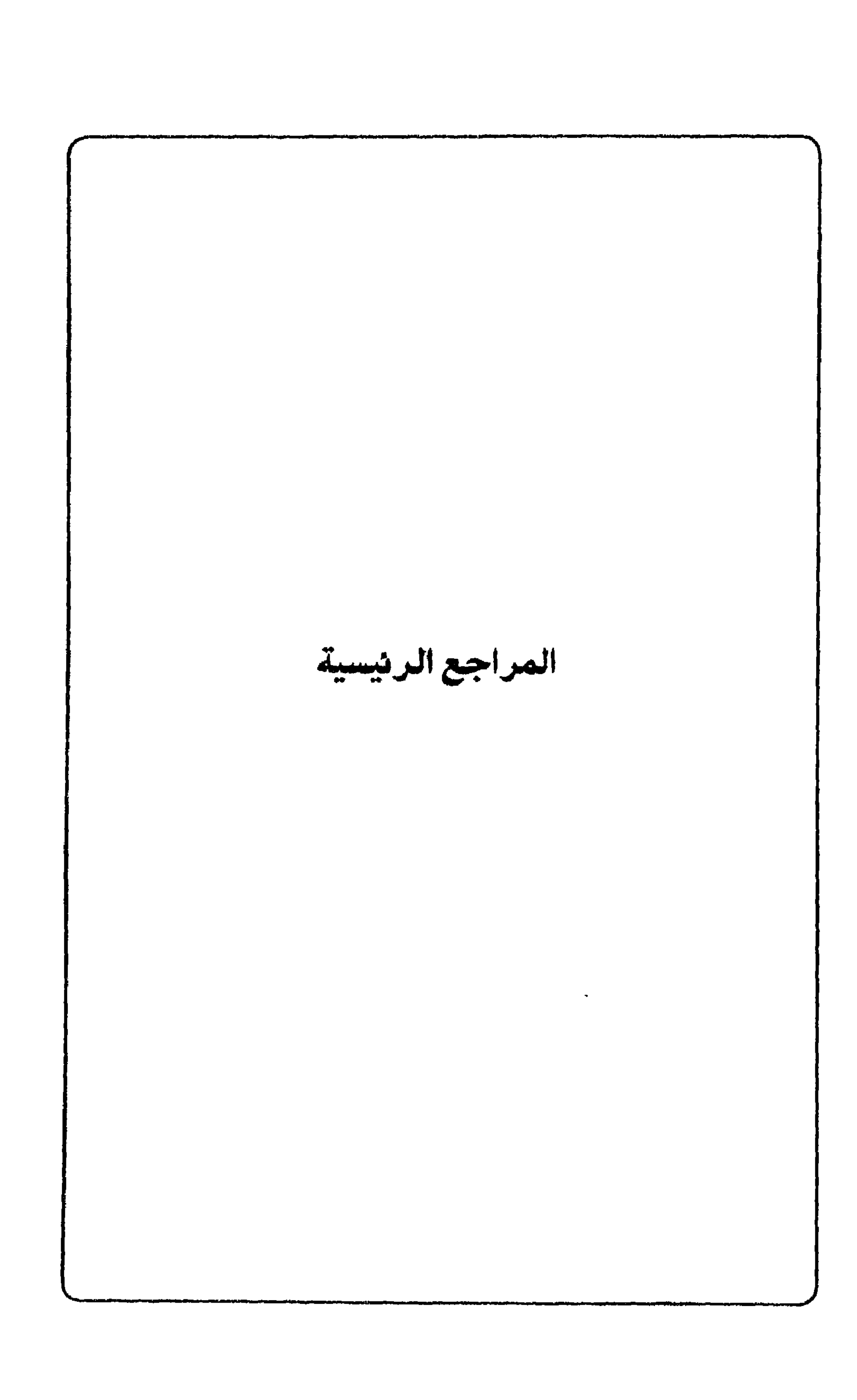
صفحة

صفحة



فهرس الأشكال

Converted by Tiff Combine - (no stamps are applied by registered version)



المراجع الرئيسية



أولاً: المراجع العربية:

- أحمد عباده سرحان (١٩٦٨): أسس الإحصاء، دار الكتب الجامعية، القاهرة.
- أحمد عباده سرحان (١٩٧١): طرق التحليل الإحصائى، الهيئة العامة للكتاب، القاهرة.
- أحمد عباده سرحان، صلاح الدين طلبه، فاروق عبد العظيم أحمد (بدون تاريخ): الاحصاء، مؤسسة شباب الجامعة، الإسكندرية.
- أحمد عباده سرحان وآخرون (١٩٦٨): تحليل الانحدار والارتباط، مكتبة عين شمس، القاهرة.
- اسماعيل محمد هاشم (١٩٦٨): الاحصاءات التطبيقية، دار المعارف، القاهرة.
- السيد سعد قاسم ولطفى هندى (١٩٦٧): مبادئ الاحصاء التجريبى، الطبعة الثانية، دار المعارف، القاهرة.
- السيد محمد خيرى (١٩٧٠): الاحصاء في البحوث النفسية والتربوية والاجتماعية، الطبعة الرابعة، دار النهضة العربية، القاهرة.
- أنيس رمسيس منصور وزكى محمد عبد الرحمن (١٩٦٧): مقدمة إلى الإحصاء، مكتبة الانجلو المصرية، القاهرة.
- بدر الدين المصرى (١٩٧٠): مذكرات في الاحصاء، دار الجامعات المصرية، الإسكندرية.
- عبد الإله أبو عياش (١٩٧٨): الإحصاء والكمبيوتر فى معالجة البيانات – مع تطبيقات جغرافية – وكالة المطبوعات، الكويت.
- عبد الرحمن البدرى (١٩٦١): مبادئ الطرق الاحصائية، دار النهضة العربية، القاهرة.
- عبد المجيد فراج (١٩٧٠): الأسلوب الاحصائى، الطبعة الثانية، مكتبة القاهرة الحديثة، القاهرة.
- محمد الفرا (١٩٧٥): مناهج البحث فى الجغرافيا بالوسائل الكمية، الطبعة الثانية، وكالة المطبوعات، الكويت.
- محمد على بشر ومحمد الروبى (١٩٨٠): مقدمة فى طرق الاحصاء وتصميم التجارب، دار المطبوعات الجديدة، الإسكندرية.
- محمد مظلوم حمدى (١٩٦١): طرق الاحصاء، دار المعارف، القاهرة.
- ناصر عبد الله الصالح ومحمد محمود السريانى (١٩٧٩): الجغرافية الكمية والاحصائية – أسس وتطبيقات – مكة المكرمة، المملكة العربية السعودية.

ثانيا: المراجع الأجنبية:

- Alder, H.L. & Roesseler, E.B. (1964): Introduction to Probility and Statistics, W.H. Freeman.
- Berry, B. & Baker, A. (1968): "Geographyic Sampling" in Berry & Marble, Spatial Analysis, Englewood Cliffs, Prentice - Hall Inc., N.J., pp. 91 - 160.
- Burton Ian (1963): "The Quantitative Revolution and Theoretical Geography" The Canadian Geography.
- Cole, J. P. & King, C.A.M. (1970): Quantitative Geography Techniques and Theories in Geography, 3rd Ed., wiley. London.
- Davis, J.C. (1973): Statistics and Data Analysis in Geology, wiley, New York.
- Davis, W.K. (1972): The Conceptual Revolution in Geography, Univ. of London Press, London.
- Dorrnkawp, J. C. & King, C.A.M. (1971): Numerical Analysis in Geomorphology, Arnold, London.
- Ebdon, D. (1977): Statistics in Geography, Basil, Blackwell, Oxford.
- Gerogry, S. (1970); Statistical Methods and Geographers, Longman, London.
- Hammond, R. & Mc Cullagh, P.S. (1974) Quantitative Techniques in Geography, Clarendon Press, Oxford.
- Harper, W. M. (1974): Statistics, 4th ed., M. & E. Handbook, London.
- Harvey, D. W. (1969): An Qutline of Statistics, Longman, London.
- Kendall, M.G. (1970) Rank Correlation Methods, Griffin, London.
- King, L. J. 91967): Statistical Analysis in Geography, Englewood Cliffs, N. J.
- Koch, G.S. & Link, R.F. (1971): Statistical Analysis of Geological Data, vol. II, Wiley, New York.

- Lindly, D.V. & Miller, J.C.P. (1953): Cambridge Elementary Statistical Tables, Cambridge.
- Morkhouse, F.J. & Wilkinson, H. R., (1971): Maps and Diagrams, Methuen, London..
- Moroney, M.J. (1975): Facts from Figures, Penguin Book, London.
- Siegle, S., (1956): Non-parametric Statistical for the Behavioural Sciences, Mc Graw-Hill, New York.
- Spiegle, M.R. (1972): Theory and Problems of Statistics, McGraw-Hill, New York.
- Stoodley, K.D.C. (1974): Basic Statistical Techniques, Bradford Univ. Press,
- Till, R. (1974): Statistical Methods for Earth Scientists, MacMillan Press, Ltd., London.
- Toyne, P. & Newby, P.T., (1971) Techniques in Human Geography MacMillan Press, Ltd., London.
- Von Mises, R. (1964): Mathematical Theory of Probability and Statistics, Academic Press, New York.
- Yeates, Maurice (1974): An Introduction to Quantitative Analysis in Human Geography, McGraw-Hill, New York.
- Yevjevich, V.L. (1971): Probability and Statistics in Hydrology, Water Resources publication, Colorado.
- Yule, G.U. & Kendaly, M.G. (1953): An Introduction to the Theory of Statistics, Griffin, London.



الجداول

جدول رقم (م - ٢)

في شرح جداول اللوغاريتمات وكيفية استعمالها

٧ جداول اللوغاريتمات العادية محسوبة على مقتضى الأساس ١٠ وأول من حسب هذه اللوغاريتمات هنري بريجز (Henry Briggs) سنة ١٦١٥ ميلادية بناء على توصية نبير (Napier, له ويقال للوغاريتمات المحسوبة على هذا الأساس اللوغاريتمات العادية أو اللوغاريتمات البريجزية (نسبة إلى الرجل بريجز الذي أدخلها).

٨ في هذه الجداول تكون لوغاريتمات الأعداد التي هي قوى للعدد ١٠ أعداداً صحيحة فمثلا

$\log_{10} 1000 = 3$ لأن $10^{3} = 1000$	$\log_{10} 0.1 = -1$ لأن $10^{-1} = \frac{1}{10}$
$\log_{10} 100 = 2$ « $10^{2} = 100$	$\log_{10} 0.01 = -2$ « $10^{-2} = \frac{1}{100}$
$\log_{10} 10 = 1$ « $10^{1} = 10$	$\log_{10} 0.001 = -3$ « $10^{-3} = \frac{1}{1000}$
$\log_{10} 1 = 0$ « $10^{0} = 1$	$\log_{10} 0.0001 = -4$ « $10^{-4} = \frac{1}{10000}$

٩ لوغاريتمات الأعداد التي ليست قوى للعدد ١٠ تتركب من عدد صحيح ومن كسر عشري ويقال للعدد الصحيح العدد البياني وللكسر الجزء العشري.

١٠ العدد البياني من لوغاريتم أي عدد أكبر من الواحد يكون موجباً ويساوي عدد أرقامه الصحيحة ناقصاً واحداً.

فالجزء البياني من لوغاريتم ٦٣٤٥٠ هو ٤
« « « ٦٣٤,٥ هو ٢
« « « ٦,٣٤٥ هو ٠

١١ العدد البياني من لوغاريتم أي عدد أصغر من الواحد يكون سالباً ويساوي عدد الأصفار التي تلي الشرطة العشرية مباشرة مضافاً إليه واحد.

فالجزء البياني من لوغاريتم ٠,٦٣٤٥ هو $\bar{1}$
« « « ٠,٠٠٦٣٤٥ هو $\bar{3}$
« « « ٠,٠٠٠٠٦٣٤٥ هو $\bar{5}$

(تنبيه) عندما يكون العدد البياني سالباً تكتب العلامة (ـــ) فوق العدد البياني مثل $\bar{2}$

١٢ الأعداد المركبة من أرقام متحدة ذات ترتيب واحد بلا تختلف إلا موضع العلامة العشرية تكون لوغاريتماتها متحدة فى الجزء العشرى ومختلفة فى العدد البيانى .

فالأجزاء العشرية من لوغاريتمات الأعداد ٦٣٢٥٠ ، ٦٣,٢٥ ، ٠,٠٦٣٢٥ كلها متساوية

١٣ لإيجاد الجزء العشرى من لوغاريتم عدد لا تزيد أرقامه المعنوية على رقم واحد مثل ٨ أو ٨٠٠ أو ٨،٠

	٠	١	٢	٣	٤	٥	٦	٧	٨	٩	الفروق								
											١	٢	٣	٤	٥	٦	٧	٨	٩
٨٠	٩٠٣١	٩٠٣٦	٩٠٤٢	٩٠٤٧	٩٠٥٣	٩٠٥٨	٩٠٦٣	٩٠٦٩	٩٠٧٤	٩٠٧٩	١	١	٢	٢	٣	٣	٤	٤	٥

نبحث عن العدد ٨٠ فى صفحات جدول لوغاريتمات الأعداد فى الصف الرأسى الأول ونبحث عن (٠) فى الصف الأفقى الأول من هذه الصفحة ثم نتبع الصف الأفقى المبدوء بالعدد ٨٠ والصف الرأسى المبدوء بصفر فنجد فى متقاطع هذين الصفين العدد ٩٠٣١ فيكون هو الجزء العشرى من لوغاريتم ٨ أو ٨٠٠ أو ٨،٠

١٤ لإيجاد الجزء العشرى من لوغاريتم عدد مركب من رقمين معنويين مثل ٨٥ أو ٨٥٠ أو ٨,٥

	٠	١	٢	٣	٤	٥	٦	٧	٨	٩	الفروق								
											١	٢	٣	٤	٥	٦	٧	٨	٩
٨٥	٩٢٩٤	٩٢٩٩	٩٣٠٤	٩٣٠٩	٩٣١٥	٩٣٢٠	٩٣٢٥	٩٣٣٠	٩٣٣٥	٩٣٤٠	١	١	٢	٢	٣	٣	٤	٤	٥

نبحث عن العدد ٨٥ فى صفحات جدول اللوغاريتمات فى الصف الرأسى الأول ونبحث عن (٠) فى الصف الأفقى الأول من هذه الصفحة ثم نتبع الصف الأفقى المبدوء بالعدد ٨٥ والصف الرأسى المبدوء بصفر فنجد فى متقاطع هذين الصفين ٩٢٩٤ فيكون هو الجزء العشرى من لوغاريتم ٨٥ أو ٨٥٠ أو ٨،٥ .

١٥ لإيجاد الجزء العشرى من لوغاريتم عدد مركب من ثلاثة أرقام معنوية مثل ٨٥٦ أو ٨,٥٦ أو ٨٥٦٠٠

نبحث عن العدد المركب من الرقمين الأولين من يسار العدد (وهو ٨٥) فى صفحات الجدول فى الصف الرأسى الأول ونبحث عن الرقم الثالث ٦ فى الصف الأفقى الأول من هذه الصفحة ثم نتبع الصف الأفقى المبدوء بالعدد ٨٥ والصف الرأسى المبدوء برقم ٦ فنجد فى متقاطع هذين الصفين ٩٣٢٥ فيكون هو الجزء العشرى من لوغاريتم ٨٥٦ أو ٨,٥٦ أو ٨٥٦٠٠

١٦ لإيجاد الجزء العشرى من لوغاريتم عدد مركب من أربعة أرقام معنوية مثل ٨٥٦٢ أو ٨٥,٦٢ أو ٨٥٦٢٠

نبحث عن الجزء العشرى للعدد ٨٥٦ بالطريقة السابقة فنجد أنه ٩٣٢٥ ثم نبحث عن ٢ فى الصف الأفقى الأول من أعمدة الفروق ونتتبع الصف الأفقى المبدوء بالعدد ٨٥ والصف الرأسى المبدوء بالفرق ٢ فنجد فى متقاطع هذين الصفين ١ فيكون هو العدد الذى يلزم إضافته إلى ٩٣٢٥ لينتج الجزء العشرى من لوغاريتم العدد ٨٥٦٢ .

وعلى ذلك يكون ٩٣٢٥ + ١ (٠,٩٣٢٦) هو الجزء العشرى من لوغاريتم العدد ٨٥٦٢ أو ٨٥,٦٢ أو ٨٥٦٢٠ .

١٧ مما تقدم نعلم طريقة إيجاد لوغاريتم أى عدد لا يزيد على أربعة أرقام وذلك بأن نأتى أولاً بجزئه العشرى ثم نضيف إليه عدده البيانى .

فمثلا لو ٦٣٤٥٠ = ٤,٨٠٢٤

لو ٦٠٣٤٥ = ٠,٨٠٢٤

لو ٠,٦٣٤٥ = ١,٨٠٤٢

لو ٠,٠٠٦٣٤٥ = ٣٠٨٠٤٢

١٨ لإيجاد العدد المقابل للوغاريتم معلوم :

نبحث بطريقة مماثلة للطريقة المبينة ببند ١٦ عن العدد المقابل للجزء العشرى من هذا اللوغاريتم فى جدول الأعداد المقابلة للوغاريتمات ثم نعدل هذا العدد كما يقتضيه العدد البيانى للوغاريتم بأن نضع على يمينه أصفاراً أو نفصل منه أرقاماً عشرية .

فإذا أردنا إيجاد العدد الذى لوغاريتمه هو ٢,٠٦٧٤ نجرى العمل هكذا .

	٠	١	٢	٣	٤	٥	٦	٧	٨	٩	الفروق ١٢٣	٤٥٦	٧٨٩
٠,٠٦	١١٤٨	١١٥١	١١٥٣	١١٥٦	١١٥٩	١١٦١	١١٦٤	١١٦٧	١١٦٩	١١٧٢	٠١١	١١٢	٢٢٢

نبحث عن ٠,٠٦ فى صفحات جدول الأعداد المقابل للوغاريتمات فى الصف الرأسى الأول ونبحث عن الرقم العشرى الثالث ٧ فى الصف الأفقى الأول من هذه الصفحة ثم نتتبع الصف الأفقى المبدوء بالعدد ٠,٠٦ والصف الرأسى المبدوء بالعدد ٧ فنجد فى متقاطع هذين الصفين ١١٦٧ ثم نبحث عن الرقم العشرى الرابع ٤ فى الصف الأفقى الأول من أعمدة الفروق ونتتبع الصف الأفقى المبدوء بالعدد ٠,٠٦ والصف الرأسى المبدوء بالفرق ٤ فنجد فى متقاطع هذين الصفين ١ فيكون هو العدد الذى يلزم إضافته إلى ١١٦٧ لتنتج الأرقام المكونة للعدد المطلوب .

وعلى ذلك تكون أرقام العدد الجارى البحث عنه تساوى ١١٦٧ + ١ (أى ١١٦٨) .

وبما أن العدد البيانى للوغاريتم هو ٢ يكون عدد أرقامه الصحيحة ٣ ويكون العدد الذى لوغاريتمه ٢,٠٦٧٤ هو ١١٦,٨ .

جدول رقم (م – ٣)

جدول لوغاريتمات الأعداد

العدد	٠	١	٢	٣	٤	٥	٦	٧	٨	٩	الفروق ١	٢	٣	٤	٥	٦	٧	٨	٩
١٠	٠	٠٠٤٣	٠٠٨٦	٠١٢٨	٠١٧٠	٠٢١٢	٢٥٣	٠٢٩٤	٠٣٣٤	٠٣٧٤	٤	٨	١٢	١٧	٢١	٢٥	٢٩	٣٣	٣٧
١١	٤١٤	٤٥٣	٤٩٢	٠٥٣١	٥٦٩	٦٠٧	٠٦٤٥	٦٨٢	٧١٩	٧٥٥	٤	٨	١١	١٥	١٩	٢٣	٢٦	٣٠	٣٤
١٢	٠٧٩٢	٠٨٢٨	٠٨٦٤	٠٨٩٩	٠٩٣٤	٠٩٦٩	١٠٠٤	١٠٣٨	١٠٧٢	١١٠٦	٣	٧	١٠	١٤	١٧	٢١	٢٤	٢٨	٣١
١٣	١١٣٩	١١٧٣	١٢٠٦	١٢٣٩	١٢٧١	١٣٠٣	١٣٣٥	١٣٦٧	١٣٩٩	١٤٣٠	٣	٦	١٠	١٣	١٦	١٩	٢٣	٢٦	٢٩
١٤	١٤٦١	١٤٩٢	١٥٢٣	١٥٥٣	١٥٨٤	١٦١٤	١٦٤٤	١٦٧٣	١٧٠٣	١٧٣٢	٣	٦	٩	١٢	١٥	١٨	٢١	٢٤	٢٧
١٥	١٧٦١	١٧٩٠	١٨١٨	١٨٤٧	١٨٧٥	١٩٠٣	١٩٣١	١٩٥٩	١٩٨٧	٢٠١٤	٣	٦	٨	١١	١٤	١٧	٢٠	٢٢	٢٥
١٦	٢٠٤١	٢٠٦٨	٢٠٩٥	٢١٢٢	٢١٤٨	٢١٧٥	٢٢٠١	٢٢٢٧	٢٢٥٣	٢٢٧٩	٣	٥	٨	١١	١٣	١٦	١٨	٢١	٢٤
١٧	٢٣٠٤	٢٣٣٠	٢٣٥٥	٢٣٨٠	٢٤٠٥	٢٤٣٠	٢٤٥٥	٢٤٨٠	٢٥٠٤	٢٥٢٩	٢	٥	٧	١٠	١٢	١٥	١٧	٢٠	٢٢
١٨	٢٥٥٣	٢٥٧٧	٢٦٠١	٢٦٢٥	٢٦٤٨	٢٦٧٢	٢٦٩٥	٢٧١٨	٢٧٤٢	٢٧٦٥	٢	٥	٧	٩	١٢	١٤	١٦	١٩	٢١
١٩	٢٧٨٨	٢٨١٠	٢٨٣٣	٢٨٥٦	٢٨٧٨	٢٩٠٠	٢٩٢٣	٢٩٤٥	٢٩٦٧	٢٩٨٩	٢	٤	٧	٩	١١	١٣	١٦	١٨	٢٠
٢٠	٣٠١٠	٣٠٣٢	٣٠٥٤	٣٠٧٥	٣٠٩٦	٣١١٨	٣١٣٩	٣١٦٠	٣١٨١	٣٢٠١	٢	٤	٦	٨	١١	١٣	١٥	١٧	١٩
٢١	٣٢٢٢	٣٢٤٣	٣٢٦٣	٣٢٨٤	٣٣٠٤	٣٣٢٤	٣٣٤٥	٣٣٦٥	٣٣٨٥	٣٤٠٤	٢	٤	٦	٨	١٠	١٢	١٤	١٦	١٨
٢٢	٣٤٢٤	٣٤٤٤	٣٤٦٤	٣٤٨٣	٣٥٠٢	٣٥٢٢	٣٥٤١	٣٥٦٠	٣٥٧٩	٣٥٩٨	٢	٤	٦	٨	١٠	١٢	١٤	١٥	١٧
٢٣	٣٦١٧	٣٦٣٦	٣٦٥٥	٣٦٧٤	٣٦٩٢	٣٧١١	٣٧٢٩	٣٧٤٧	٣٧٦٦	٣٧٨٤	٢	٤	٦	٧	٩	١١	١٣	١٥	١٧
٢٤	٣٨٠٢	٣٨٢٠	٣٨٣٨	٣٨٥٦	٣٨٧٤	٣٨٩٢	٣٩٠٩	٣٩٢٧	٣٩٤٥	٣٩٦٢	٢	٤	٥	٧	٩	١١	١٢	١٤	١٦
٢٥	٣٩٧٩	٣٩٩٧	٤٠١٤	٤٠٣١	٤٠٤٨	٤٠٦٥	٤٠٨٢	٤٠٩٩	٤١١٦	٤١٣٣	٢	٣	٥	٧	٩	١٠	١٢	١٤	١٥
٢٦	٤١٥٠	٤١٦٦	٤١٨٣	٤٢٠٠	٤٢١٦	٤٢٣٢	٤٢٤٩	٤٢٦٥	٤٢٨١	٤٢٩٨	٢	٣	٥	٧	٨	١٠	١١	١٣	١٥
٢٧	٤٣١٤	٤٣٣٠	٤٣٤٦	٤٣٦٢	٤٣٧٨	٤٣٩٣	٤٤٠٩	٤٤٢٥	٤٤٤٠	٤٤٥٦	٢	٣	٥	٦	٨	٩	١١	١٣	١٤
٢٨	٤٤٧٢	٤٤٨٧	٤٥٠٢	٤٥١٨	٤٥٣٣	٤٥٤٨	٤٥٦٤	٤٥٧٩	٤٥٩٤	٤٦٠٩	٢	٣	٥	٦	٨	٩	١١	١٢	١٤
٢٩	٤٦٢٤	٤٦٣٩	٤٦٥٤	٤٦٦٩	٤٦٨٣	٤٦٩٨	٤٧١٣	٤٧٢٨	٤٧٤٢	٤٧٥٧	١	٣	٤	٦	٧	٩	١٠	١٢	١٣
٣٠	٤٧٧١	٤٧٨٦	٤٨٠٠	٤٨١٤	٤٨٢٩	٤٨٤٣	٤٨٥٧	٤٨٧١	٤٨٨٦	٤٩٠٠	١	٣	٤	٦	٧	٩	١٠	١١	١٣
٣١	٤٩١٤	٤٩٢٨	٤٩٤٢	٤٩٥٥	٤٩٦٩	٤٩٨٣	٤٩٩٧	٥٠١١	٥٠٢٤	٥٠٣٨	١	٣	٤	٦	٧	٨	١٠	١١	١٢
٣٢	٥٠٥١	٥٠٦٥	٥٠٧٩	٥٠٩٢	٥١٠٥	٥١١٩	٥١٣٢	٥١٤٥	٥١٥٩	٥١٧٢	١	٣	٤	٥	٧	٨	٩	١١	١٢
٣٣	٥١٨٥	٥١٩٨	٥٢١١	٥٢٢٤	٥٢٣٧	٥٢٥٠	٥٢٦٣	٥٢٧٦	٥٢٨٩	٥٣٠٢	١	٣	٤	٥	٦	٨	٩	١٠	١٢
٣٤	٥٣١٥	٥٣٢٨	٥٣٤٠	٥٣٥٣	٥٣٦٦	٥٣٧٨	٥٣٩١	٥٤٠٣	٥٤١٦	٥٤٢٨	١	٣	٤	٥	٦	٨	٩	١٠	١١
٣٥	٥٤٤١	٥٤٥٣	٥٤٦٥	٥٤٧٨	٥٤٩٠	٥٥٠٢	٥٥١٤	٥٥٢٧	٥٥٣٩	٥٥٥١	١	٢	٤	٥	٦	٧	٩	١٠	١١
٣٦	٥٥٦٣	٥٥٧٥	٥٥٨٧	٥٥٩٩	٥٦١١	٥٦٢٣	٥٦٣٥	٥٦٤٧	٥٦٥٨	٥٦٧٠	١	٢	٤	٥	٦	٧	٨	١٠	١١
٣٧	٥٦٨٢	٥٦٩٤	٥٧٠٥	٥٧١٧	٥٧٢٩	٥٧٤٠	٥٧٥٢	٥٧٦٣	٥٧٧٥	٥٧٨٦	١	٢	٣	٥	٦	٧	٨	٩	١٠
٣٨	٥٧٩٨	٥٨٠٩	٥٨٢١	٥٨٣٢	٥٨٤٣	٥٨٥٥	٥٨٦٦	٥٨٧٧	٥٨٨٨	٥٨٩٩	١	٢	٣	٥	٦	٧	٨	٩	١٠
٣٩	٥٩١١	٥٩٢٢	٥٩٣٣	٥٩٤٤	٥٩٥٥	٥٩٦٦	٥٩٧٧	٥٩٨٨	٥٩٩٩	٦٠١٠	١	٢	٣	٤	٥	٧	٨	٩	١٠
٤٠	٦٠٢١	٦٠٣١	٦٠٤٢	٦٠٥٣	٦٠٦٤	٦٠٧٥	٦٠٨٥	٦٠٩٦	٦١٠٧	٦١١٧	١	٢	٣	٤	٥	٦	٨	٩	١٠
٤١	٦١٢٨	٦١٣٨	٦١٤٩	٦١٦٠	٦١٧٠	٦١٨٠	٦١٩١	٦٢٠١	٦٢١٢	٦٢٢٢	١	٢	٣	٤	٥	٦	٧	٨	٩
٤٢	٦٢٣٢	٦٢٤٣	٦٢٥٣	٦٢٦٣	٦٢٧٤	٦٢٨٤	٦٢٩٤	٦٣٠٤	٦٣١٤	٦٣٢٥	١	٢	٣	٤	٥	٦	٧	٨	٩
٤٣	٦٣٣٥	٦٣٤٥	٦٣٥٥	٦٣٦٥	٦٣٧٥	٦٣٨٥	٦٣٩٥	٦٤٠٥	٦٤١٥	٦٤٢٥	١	٢	٣	٤	٥	٦	٧	٨	٩
٤٤	٦٤٣٥	٦٤٤٤	٦٤٥٤	٦٤٦٤	٦٤٧٤	٦٤٨٤	٦٤٩٣	٦٥٠٣	٦٥١٣	٦٥٢٢	١	٢	٣	٤	٥	٦	٧	٨	٩
٤٥	٦٥٣٢	٦٥٤٢	٦٥٥١	٦٥٦١	٦٥٧١	٦٥٨٠	٦٥٩٠	٦٥٩٩	٦٦٠٩	٦٦١٨	١	٢	٣	٤	٥	٦	٧	٨	٩
٤٦	٦٦٢٨	٦٦٣٧	٦٦٤٦	٦٦٥٦	٦٦٦٥	٦٦٧٥	٦٦٨٤	٦٦٩٣	٦٧٠٢	٦٧١٢	١	٢	٣	٤	٥	٦	٧	٧	٨
٤٧	٦٧٢١	٦٧٣٠	٦٧٣٩	٦٧٤٩	٦٧٥٨	٦٧٦٧	٦٧٧٦	٦٧٨٥	٦٧٩٤	٦٨٠٣	١	٢	٣	٤	٥	٥	٦	٧	٨
٤٨	٦٨١٢	٦٨٢١	٦٨٣٠	٦٨٣٩	٦٨٤٨	٦٨٥٧	٦٨٦٦	٦٨٧٥	٦٨٨٤	٦٨٩٣	١	٢	٣	٤	٤	٥	٦	٧	٨
٤٩	٦٩٠٢	٦٩١١	٦٩٢٠	٦٩٢٨	٦٩٣٧	٦٩٤٦	٦٩٥٥	٦٩٦٤	٦٩٧٢	٦٩٨١	١	٢	٣	٤	٤	٥	٦	٧	٨
٥٠	٦٩٩٠	٦٩٩٨	٧٠٠٧	٧٠١٦	٧٠٢٤	٧٠٣٣	٧٠٤٢	٧٠٥٠	٧٠٥٩	٧٠٦٧	١	٢	٣	٣	٤	٥	٦	٧	٨
٥١	٧٠٧٦	٧٠٨٤	٧٠٩٣	٧١٠١	٧١١٠	٧١١٨	٧١٢٦	٧١٣٥	٧١٤٣	٧١٥٢	١	٢	٣	٣	٤	٥	٦	٧	٨
٥٢	٧١٦٠	٧١٦٨	٧١٧٧	٧١٨٥	٧١٩٣	٧٢٠٢	٧٢١٠	٧٢١٨	٧٢٢٦	٧٢٣٥	١	٢	٢	٣	٤	٥	٦	٧	٧
٥٣	٧٢٤٣	٧٢٥١	٧٢٥٩	٧٢٦٧	٧٢٧٥	٧٢٨٤	٧٢٩٢	٧٣٠٠	٧٣٠٨	٧٣١٦	١	٢	٢	٣	٤	٥	٦	٦	٧
٥٤	٧٣٢٤	٧٣٣٢	٧٣٤٠	٧٣٤٨	٧٣٥٦	٧٣٦٤	٧٣٧٢	٧٣٨٠	٧٣٨٨	٧٣٩٦	١	٢	٢	٣	٤	٥	٦	٦	٧

(تابع) جدول لوغاريتمات الأعداد

العدد	٠	١	٢	٣	٤	٥	٦	٧	٨	٩	الفروق ١	٢	٣	٤	٥	٦	٧	٨	٩
٥٥	٧٤٠٤	٧٤١٢	٧٤١٩	٧٤٢٧	٧٤٣٥	٧٤٤٣	٧٤٥١	٧٤٥٩	٧٤٦٦	٧٤٧٤	١	٢	٢	٣	٤	٥	٥	٦	٧
٥٦	٧٤٨٢	٧٤٩٠	٧٤٩٧	٧٥٠٥	٧٥١٣	٧٥٢٠	٧٥٢٨	٧٥٣٦	٧٥٤٣	٧٥٥١	١	٢	٢	٣	٤	٥	٥	٦	٧
٥٧	٧٥٥٩	٧٥٦٦	٧٥٧٤	٧٥٨٢	٧٥٨٩	٧٥٩٧	٧٦٠٤	٧٦١٢	٧٦١٩	٧٦٢٧	١	٢	٢	٣	٤	٥	٥	٦	٧
٥٨	٧٦٣٤	٧٦٤٢	٧٦٤٩	٧٦٥٧	٧٦٦٤	٧٦٧٢	٧٦٧٩	٧٦٨٦	٧٦٩٤	٧٧٠١	١	١	٢	٣	٤	٤	٥	٦	٧
٥٩	٧٧٠٩	٧٧١٦	٧٧٢٣	٧٧٣١	٧٧٣٨	٧٧٤٥	٧٧٥٢	٧٧٦٠	٧٧٦٧	٧٧٧٤	١	١	٢	٣	٤	٤	٥	٦	٧
٦٠	٧٧٨٢	٧٧٨٩	٧٧٩٦	٧٨٠٣	٧٨١٠	٧٨١٨	٧٨٢٥	٧٨٣٢	٧٨٣٩	٧٨٤٦	١	١	٢	٣	٤	٤	٥	٦	٦
٦١	٧٨٥٣	٧٨٦٠	٧٨٦٨	٧٨٧٥	٧٨٨٢	٧٨٨٩	٧٨٩٦	٧٩٠٣	٧٩١٠	٧٩١٧	١	١	٢	٣	٤	٤	٥	٦	٦
٦٢	٧٩٢٤	٧٩٣١	٧٩٣٨	٧٩٤٥	٧٩٥٢	٧٩٥٩	٧٩٦٦	٧٩٧٣	٧٩٨٠	٧٩٨٧	١	١	٢	٣	٣	٤	٥	٦	٦
٦٣	٧٩٩٣	٨٠٠٠	٨٠٠٧	٨٠١٤	٨٠٢١	٨٠٢٨	٨٠٣٥	٨٠٤١	٨٠٤٨	٨٠٥٥	١	١	٢	٣	٣	٤	٥	٥	٦
٦٤	٨٠٦٢	٨٠٦٩	٨٠٧٥	٨٠٨٢	٨٠٨٩	٨٠٩٦	٨١٠٢	٨١٠٩	٨١١٦	٨١٢٢	١	١	٢	٣	٣	٤	٥	٥	٦
٦٥	٨١٢٩	٨١٣٦	٨١٤٢	٨١٤٩	٨١٥٦	٨١٦٢	٨١٦٩	٨١٧٦	٨١٨٢	٨١٨٩	١	١	٢	٣	٣	٤	٥	٥	٦
٦٦	٨١٩٥	٨٢٠٢	٨٢٠٩	٨٢١٥	٨٢٢٢	٨٢٢٨	٨٢٣٥	٨٢٤١	٨٢٤٨	٨٢٥٤	١	١	٢	٣	٣	٤	٥	٥	٦
٦٧	٨٢٦١	٨٢٦٧	٨٢٧٤	٨٢٨٠	٨٢٨٧	٨٢٩٣	٨٢٩٩	٨٣٠٦	٨٣١٢	٨٣١٩	١	١	٢	٣	٣	٤	٥	٥	٦
٦٨	٨٣٢٥	٨٣٣١	٨٣٣٨	٨٣٤٤	٨٣٥١	٨٣٥٧	٨٣٦٣	٨٣٧٠	٨٣٧٦	٨٣٨٢	١	١	٢	٣	٣	٤	٤	٥	٦
٦٩	٨٣٨٨	٨٣٩٥	٨٤٠١	٨٤٠٧	٨٤١٤	٨٤٢٠	٨٤٢٦	٨٤٣٢	٨٤٣٩	٨٤٤٥	١	١	٢	٢	٣	٤	٤	٥	٦
٧٠	٨٤٥١	٨٤٥٧	٨٤٦٣	٨٤٧٠	٨٤٧٦	٨٤٨٢	٨٤٨٨	٨٤٩٤	٨٥٠٠	٨٥٠٦	١	١	٢	٢	٣	٤	٤	٥	٦
٧١	٨٥١٣	٨٥١٩	٨٥٢٥	٨٥٣١	٨٥٣٧	٨٥٤٣	٨٥٤٩	٨٥٥٥	٨٥٦١	٨٥٦٧	١	١	٢	٢	٣	٤	٤	٥	٥
٧٢	٨٥٧٣	٨٥٧٩	٨٥٨٥	٨٥٩١	٨٥٩٧	٨٦٠٣	٨٦٠٩	٨٦١٥	٨٦٢١	٨٦٢٧	١	١	٢	٢	٣	٤	٤	٥	٥
٧٣	٨٦٣٣	٨٦٣٩	٨٦٤٥	٨٦٥١	٨٦٥٧	٨٦٦٣	٨٦٦٩	٨٦٧٥	٨٦٨١	٨٦٨٦	١	١	٢	٢	٣	٤	٤	٥	٥
٧٤	٨٦٩٢	٨٦٩٨	٨٧٠٤	٨٧١٠	٨٧١٦	٨٧٢٢	٨٧٢٧	٨٧٣٣	٨٧٣٩	٨٧٤٥	١	١	٢	٢	٣	٤	٤	٥	٥
٧٥	٨٧٥١	٨٧٥٦	٨٧٦٢	٨٧٦٨	٨٧٧٤	٨٧٧٩	٨٧٨٥	٨٧٩١	٨٧٩٧	٨٨٠٢	١	١	٢	٢	٣	٣	٤	٥	٥
٧٦	٨٨٠٨	٨٨١٤	٨٨٢٠	٨٨٢٥	٨٨٣١	٨٨٣٧	٨٨٤٢	٨٨٤٨	٨٨٥٤	٨٨٥٩	١	١	٢	٢	٣	٣	٤	٥	٥
٧٧	٨٨٦٥	٨٨٧١	٨٨٧٦	٨٨٨٢	٨٨٨٧	٨٨٩٣	٨٨٩٩	٨٩٠٤	٨٩١٠	٨٩١٥	١	١	٢	٢	٣	٣	٤	٤	٥
٧٨	٨٩٢١	٨٩٢٧	٨٩٣٢	٨٩٣٨	٨٩٤٣	٨٩٤٩	٨٩٥٤	٨٩٦٠	٨٩٦٥	٨٩٧١	١	١	٢	٢	٣	٣	٤	٤	٥
٧٩	٨٩٧٦	٨٩٨٢	٨٩٨٧	٨٩٩٣	٨٩٩٨	٩٠٠٤	٩٠٠٩	٩٠١٥	٩٠٢٠	٩٠٢٥	١	١	٢	٢	٣	٣	٤	٤	٥
٨٠	٩٠٣١	٩٠٣٦	٩٠٤٢	٩٠٤٧	٩٠٥٣	٩٠٥٨	٩٠٦٣	٩٠٦٩	٩٠٧٤	٩٠٧٩	١	١	٢	٢	٣	٣	٤	٤	٥
٨١	٩٠٨٥	٩٠٩٠	٩٠٩٦	٩١٠١	٩١٠٦	٩١١٢	٩١١٧	٩١٢٢	٩١٢٨	٩١٣٣	١	١	٢	٢	٣	٣	٤	٤	٥
٨٢	٩١٣٨	٩١٤٣	٩١٤٩	٩١٥٤	٩١٥٩	٩١٦٥	٩١٧٠	٩١٧٥	٩١٨٠	٩١٨٦	١	١	٢	٢	٣	٣	٤	٤	٥
٨٣	٩١٩١	٩١٩٦	٩٢٠١	٩٢٠٦	٩٢١٢	٩٢١٧	٩٢٢٢	٩٢٢٧	٩٢٣٢	٩٢٣٨	١	١	٢	٢	٣	٣	٤	٤	٥
٨٤	٩٢٤٣	٩٢٤٨	٩٢٥٣	٩٢٥٨	٩٢٦٣	٩٢٦٩	٩٢٧٤	٩٢٧٩	٩٢٨٤	٩٢٨٩	١	١	٢	٢	٣	٣	٤	٤	٥
٨٥	٩٢٩٤	٩٢٩٩	٩٣٠٤	٩٣٠٩	٩٣١٥	٩٣٢٠	٩٣٢٥	٩٣٣٠	٩٣٣٥	٩٣٤٠	١	١	٢	٢	٣	٣	٤	٤	٥
٨٦	٩٣٤٥	٩٣٥٠	٩٣٥٥	٩٣٦٠	٩٣٦٥	٩٣٧٠	٩٣٧٥	٩٣٨٠	٩٣٨٥	٩٣٩٠	١	١	٢	٢	٣	٣	٤	٤	٥
٨٧	٩٣٩٥	٩٤٠٠	٩٤٠٥	٩٤١٠	٩٤١٥	٩٤٢٠	٩٤٢٥	٩٤٣٠	٩٤٣٥	٩٤٤٠	٠	١	١	٢	٢	٣	٣	٤	٤
٨٨	٩٤٤٥	٩٤٥٠	٩٤٥٥	٩٤٦٠	٩٤٦٥	٩٤٦٩	٩٤٧٤	٩٤٧٩	٩٤٨٤	٩٤٨٩	٠	١	١	٢	٢	٣	٣	٤	٤
٨٩	٩٤٩٤	٩٤٩٩	٩٥٠٤	٩٥٠٩	٩٥١٣	٩٥١٨	٩٥٢٣	٩٥٢٨	٩٥٣٣	٩٥٣٨	٠	١	١	٢	٢	٣	٣	٤	٤
٩٠	٩٥٤٢	٩٥٤٧	٩٥٥٢	٩٥٥٧	٩٥٦٢	٩٥٦٦	٩٥٧١	٩٥٧٦	٩٥٨١	٩٥٨٦	٠	١	١	٢	٢	٣	٣	٤	٤
٩١	٩٥٩٠	٩٥٩٥	٩٦٠٠	٩٦٠٥	٩٦٠٩	٩٦١٤	٩٦١٩	٩٦٢٤	٩٦٢٨	٩٦٣٣	٠	١	١	٢	٢	٣	٣	٤	٤
٩٢	٩٦٣٨	٩٦٤٣	٩٦٤٧	٩٦٥٢	٩٦٥٧	٩٦٦١	٩٦٦٦	٩٦٧١	٩٦٧٥	٩٦٨٠	٠	١	١	٢	٢	٣	٣	٤	٤
٩٣	٩٦٨٥	٩٦٨٩	٩٦٩٤	٩٦٩٩	٩٧٠٣	٩٧٠٨	٩٧١٣	٩٧١٧	٩٧٢٢	٩٧٢٧	٠	١	١	٢	٢	٣	٣	٤	٤
٩٤	٩٧٣١	٩٧٣٦	٩٧٤١	٩٧٤٥	٩٧٥٠	٩٧٥٤	٩٧٥٩	٩٧٦٣	٩٧٦٨	٩٧٧٣	٠	١	١	٢	٢	٣	٣	٤	٤
٩٥	٩٧٧٧	٩٧٨٢	٩٧٨٦	٩٧٩١	٩٧٩٥	٩٨٠٠	٩٨٠٥	٩٨٠٩	٩٨١٤	٩٨١٨	٠	١	١	٢	٢	٣	٣	٤	٤
٩٦	٩٨٢٣	٩٨٢٧	٩٨٣٢	٩٨٣٦	٩٨٤١	٩٨٤٥	٩٨٥٠	٩٨٥٤	٩٨٥٩	٩٨٦٣	٠	١	١	٢	٢	٣	٣	٤	٤
٩٧	٩٨٦٨	٩٨٧٢	٩٨٧٧	٩٨٨١	٩٨٨٦	٩٨٩٠	٩٨٩٤	٩٨٩٩	٩٩٠٣	٩٩٠٨	٠	١	١	٢	٢	٣	٣	٤	٤
٩٨	٩٩١٢	٩٩١٧	٩٩٢١	٩٩٢٦	٩٩٣٠	٩٩٣٤	٩٩٣٩	٩٩٤٣	٩٩٤٨	٩٩٥٢	٠	١	١	٢	٢	٣	٣	٤	٤
٩٩	٩٩٥٦	٩٩٦١	٩٩٦٥	٩٩٦٩	٩٩٧٤	٩٩٧٨	٩٩٨٣	٩٩٨٧	٩٩٩١	٩٩٩٦	٠	١	١	٢	٢	٣	٣	٣	٤

جدول رقم (٣ - ٤)

جدول الأعداد المقابلة للوغاريتمات

	٠	١	٢	٣	٤	٥	٦	٧	٨	٩	الفروق								
											١	٢	٣	٤	٥	٦	٧	٨	٩
٠٫٠٠	١٠٠٠	١٠٠٢	١٠٠٥	١٠٠٧	١٠٠٩	١٠١٢	١٠١٤	١٠١٦	١٠١٩	١٠٢١	٠	٠	١	١	١	١	٢	٢	٢
٠٫٠١	١٠٢٣	١٠٢٦	١٠٢٨	١٠٣٠	١٠٣٣	١٠٣٥	١٠٣٨	١٠٤٠	١٠٤٢	١٠٤٥	٠	٠	١	١	١	١	٢	٢	٢
٠٫٠٢	١٠٤٧	١٠٥٠	١٠٥٢	١٠٥٤	١٠٥٧	١٠٥٩	١٠٦٢	١٠٦٤	١٠٦٧	١٠٦٩	٠	٠	١	١	١	١	٢	٢	٢
٠٫٠٣	١٠٧٢	١٠٧٤	١٠٧٦	١٠٧٩	١٠٨١	١٠٨٤	١٠٨٦	١٠٨٩	١٠٩١	١٠٩٤	٠	٠	١	١	١	١	٢	٢	٢
٠٫٠٤	١٠٩٦	١٠٩٩	١١٠٢	١١٠٤	١١٠٧	١١٠٩	١١١٢	١١١٤	١١١٧	١١١٩	٠	١	١	١	١	٢	٢	٢	٢
٠٫٠٥	١١٢٢	١١٢٥	١١٢٧	١١٣٠	١١٣٢	١١٣٥	١١٣٨	١١٤٠	١١٤٣	١١٤٦	٠	١	١	١	١	٢	٢	٢	٢
٠٫٠٦	١١٤٨	١١٥١	١١٥٣	١١٥٦	١١٥٩	١١٦١	١١٦٤	١١٦٧	١١٦٩	١١٧٢	٠	١	١	١	١	٢	٢	٢	٢
٠٫٠٧	١١٧٥	١١٧٨	١١٨٠	١١٨٣	١١٨٦	١١٨٩	١١٩١	١١٩٤	١١٩٧	١١٩٩	٠	١	١	١	١	٢	٢	٢	٢
٠٫٠٨	١٢٠٢	١٢٠٥	١٢٠٨	١٢١١	١٢١٣	١٢١٦	١٢١٩	١٢٢٢	١٢٢٥	١٢٢٧	٠	١	١	١	١	٢	٢	٢	٣
٠٫٠٩	١٢٣٠	١٢٣٣	١٢٣٦	١٢٣٩	١٢٤٢	١٢٤٥	١٢٤٧	١٢٥٠	١٢٥٣	١٢٥٦	٠	١	١	١	١	٢	٢	٢	٣
٠٫١٠	١٢٥٩	١٢٦٢	١٢٦٥	١٢٦٨	١٢٧١	١٢٧٤	١٢٧٦	١٢٧٩	١٢٨٢	١٢٨٥	٠	١	١	١	١	٢	٢	٢	٣
٠٫١١	١٢٨٨	١٢٩١	١٢٩٤	١٢٩٧	١٣٠٠	١٣٠٣	١٣٠٦	١٣٠٩	١٣١٢	١٣١٥	٠	١	١	١	٢	٢	٢	٢	٣
٠٫١٢	١٣١٨	١٣٢١	١٣٢٤	١٣٢٧	١٣٣٠	١٣٣٤	١٣٣٧	١٣٤٠	١٣٤٣	١٣٤٦	٠	١	١	١	٢	٢	٢	٢	٣
٠٫١٣	١٣٤٩	١٣٥٢	١٣٥٥	١٣٥٨	١٣٦١	١٣٦٥	١٣٦٨	١٣٧١	١٣٧٤	١٣٧٧	٠	١	١	١	٢	٢	٢	٣	٣
٠٫١٤	١٣٨٠	١٣٨٤	١٣٨٧	١٣٩٠	١٣٩٣	١٣٩٦	١٤٠٠	١٤٠٣	١٤٠٦	١٤٠٩	٠	١	١	١	٢	٢	٢	٣	٣
٠٫١٥	١٤١٣	١٤١٦	١٤١٩	١٤٢٢	١٤٢٦	١٤٢٩	١٤٣٢	١٤٣٥	١٤٣٩	١٤٤٢	٠	١	١	١	٢	٢	٢	٣	٣
٠٫١٦	١٤٤٥	١٤٤٩	١٤٥٢	١٤٥٥	١٤٥٩	١٤٦٢	١٤٦٦	١٤٦٩	١٤٧٢	١٤٧٦	٠	١	١	١	٢	٢	٢	٣	٣
٠٫١٧	١٤٧٩	١٤٨٣	١٤٨٦	١٤٨٩	١٤٩٣	١٤٩٦	١٥٠٠	١٥٠٣	١٥٠٧	١٥١٠	٠	١	١	١	٢	٢	٢	٣	٣
٠٫١٨	١٥١٤	١٥١٧	١٥٢١	١٥٢٤	١٥٢٨	١٥٣١	١٥٣٥	١٥٣٨	١٥٤٢	١٥٤٥	٠	١	١	١	٢	٢	٢	٣	٣
٠٫١٩	١٥٤٩	١٥٥٢	١٥٥٦	١٥٦٠	١٥٦٣	١٥٦٧	١٥٧٠	١٥٧٤	١٥٧٨	١٥٨١	٠	١	١	١	٢	٢	٣	٣	٣
٠٫٢٠	١٥٨٥	١٥٨٩	١٥٩٢	١٥٩٦	١٦٠٠	١٦٠٣	١٦٠٧	١٦١١	١٦١٤	١٦١٨	٠	١	١	١	٢	٢	٣	٣	٣
٠٫٢١	١٦٢٢	١٦٢٦	١٦٢٩	١٦٣٣	١٦٣٧	١٦٤١	١٦٤٤	١٦٤٨	١٦٥٢	١٦٥٦	٠	١	١	٢	٢	٢	٣	٣	٣
٠٫٢٢	١٦٦٠	١٦٦٣	١٦٦٧	١٦٧١	١٦٧٥	١٦٧٩	١٦٨٣	١٦٨٧	١٦٩٠	١٦٩٤	٠	١	١	٢	٢	٢	٣	٣	٣
٠٫٢٣	١٦٩٨	١٧٠٢	١٧٠٦	١٧١٠	١٧١٤	١٧١٨	١٧٢٢	١٧٢٦	١٧٣٠	١٧٣٤	٠	١	١	٢	٢	٢	٣	٣	٤
٠٫٢٤	١٧٣٨	١٧٤٢	١٧٤٦	١٧٥٠	١٧٥٤	١٧٥٨	١٧٦٢	١٧٦٦	١٧٧٠	١٧٧٤	٠	١	١	٢	٢	٢	٣	٣	٤
٠٫٢٥	١٧٧٨	١٧٨٢	١٧٨٦	١٧٩١	١٧٩٥	١٧٩٩	١٨٠٣	١٨٠٧	١٨١١	١٨١٦	٠	١	١	٢	٢	٢	٣	٣	٤
٠٫٢٦	١٨٢٠	١٨٢٤	١٨٢٨	١٨٣٢	١٨٣٧	١٨٤١	١٨٤٥	١٨٤٩	١٨٥٤	١٨٥٨	٠	١	١	٢	٢	٣	٣	٣	٤
٠٫٢٧	١٨٦٢	١٨٦٦	١٨٧١	١٨٧٥	١٨٧٩	١٨٨٤	١٨٨٨	١٨٩٢	١٨٩٧	١٩٠١	٠	١	١	٢	٢	٣	٣	٣	٤
٠٫٢٨	١٩٠٥	١٩١٠	١٩١٤	١٩١٩	١٩٢٣	١٩٢٨	١٩٣٢	١٩٣٦	١٩٤١	١٩٤٥	٠	١	١	٢	٢	٣	٣	٤	٤
٠٫٢٩	١٩٥٠	١٩٥٤	١٩٥٩	١٩٦٣	١٩٦٨	١٩٧٢	١٩٧٧	١٩٨٢	١٩٨٦	١٩٩١	٠	١	١	٢	٢	٣	٣	٤	٤
٠٫٣٠	١٩٩٥	٢٠٠٠	٢٠٠٤	٢٠٠٩	٢٠١٤	٢٠١٨	٢٠٢٣	٢٠٢٨	٢٠٣٢	٢٠٣٧	٠	١	١	٢	٢	٣	٣	٤	٤
٠٫٣١	٢٠٤٢	٢٠٤٦	٢٠٥١	٢٠٥٦	٢٠٦١	٢٠٦٥	٢٠٧٠	٢٠٧٥	٢٠٨٠	٢٠٨٤	٠	١	١	٢	٢	٣	٣	٤	٤
٠٫٣٢	٢٠٨٩	٢٠٩٤	٢٠٩٩	٢١٠٤	٢١٠٩	٢١١٣	٢١١٨	٢١٢٣	٢١٢٨	٢١٣٣	٠	١	١	٢	٢	٣	٣	٤	٤
٠٫٣٣	٢١٣٨	٢١٤٣	٢١٤٨	٢١٥٣	٢١٥٨	٢١٦٣	٢١٦٨	٢١٧٣	٢١٧٨	٢١٨٣	٠	١	١	٢	٢	٣	٣	٤	٤
٠٫٣٤	٢١٨٨	٢١٩٣	٢١٩٨	٢٢٠٣	٢٢٠٨	٢٢١٣	٢٢١٨	٢٢٢٣	٢٢٢٨	٢٢٣٤	١	١	٢	٢	٣	٣	٤	٤	٥
٠٫٣٥	٢٢٣٩	٢٢٤٤	٢٢٤٩	٢٢٥٤	٢٢٥٩	٢٢٦٥	٢٢٧٠	٢٢٧٥	٢٢٨٠	٢٢٨٦	١	١	٢	٢	٣	٣	٤	٤	٥
٠٫٣٦	٢٢٩١	٢٢٩٦	٢٣٠١	٢٣٠٧	٢٣١٢	٢٣١٧	٢٣٢٣	٢٣٢٨	٢٣٣٣	٢٣٣٩	١	١	٢	٢	٣	٣	٤	٤	٥
٠٫٣٧	٢٣٤٤	٢٣٥٠	٢٣٥٥	٢٣٦٠	٢٣٦٦	٢٣٧١	٢٣٧٧	٢٣٨٢	٢٣٨٨	٢٣٩٣	١	١	٢	٢	٣	٣	٤	٤	٥
٠٫٣٨	٢٣٩٩	٢٤٠٤	٢٤١٠	٢٤١٥	٢٤٢١	٢٤٢٧	٢٤٣٢	٢٤٣٨	٢٤٤٣	٢٤٤٩	١	١	٢	٢	٣	٣	٤	٤	٥
٠٫٣٩	٢٤٥٥	٢٤٦٠	٢٤٦٦	٢٤٧٢	٢٤٧٧	٢٤٨٣	٢٤٨٩	٢٤٩٥	٢٥٠٠	٢٥٠٦	١	١	٢	٢	٣	٣	٤	٥	٥
٠٫٤٠	٢٥١٢	٢٥١٨	٢٥٢٣	٢٥٢٩	٢٥٣٥	٢٥٤١	٢٥٤٧	٢٥٥٣	٢٥٥٩	٢٥٦٤	١	١	٢	٢	٣	٤	٤	٥	٥
٠٫٤١	٢٥٧٠	٢٥٧٦	٢٥٨٢	٢٥٨٨	٢٥٩٤	٢٦٠٠	٢٦٠٦	٢٦١٢	٢٦١٨	٢٦٢٤	١	١	٢	٢	٣	٤	٤	٥	٥
٠٫٤٢	٢٦٣٠	٢٦٣٦	٢٦٤٢	٢٦٤٩	٢٦٥٥	٢٦٦١	٢٦٦٧	٢٦٧٣	٢٦٧٩	٢٦٨٥	١	١	٢	٢	٣	٤	٤	٥	٦
٠٫٤٣	٢٦٩٢	٢٦٩٨	٢٧٠٤	٢٧١٠	٢٧١٦	٢٧٢٣	٢٧٢٩	٢٧٣٥	٢٧٤٢	٢٧٤٨	١	١	٢	٣	٣	٤	٤	٥	٦
٠٫٤٤	٢٧٥٤	٢٧٦١	٢٧٦٧	٢٧٧٣	٢٧٨٠	٢٧٨٦	٢٧٩٣	٢٧٩٩	٢٨٠٥	٢٨١٢	١	١	٢	٣	٣	٤	٤	٥	٦
٠٫٤٥	٢٨١٨	٢٨٢٥	٢٨٣١	٢٨٣٨	٢٨٤٤	٢٨٥١	٢٨٥٨	٢٨٦٤	٢٨٧١	٢٨٧٧	١	١	٢	٣	٣	٤	٥	٥	٦
٠٫٤٦	٢٨٨٤	٢٨٩١	٢٨٩٧	٢٩٠٤	٢٩١١	٢٩١٧	٢٩٢٤	٢٩٣١	٢٩٣٨	٢٩٤٤	١	١	٢	٣	٣	٤	٥	٥	٦
٠٫٤٧	٢٩٥١	٢٩٥٨	٢٩٦٥	٢٩٧٢	٢٩٧٩	٢٩٨٥	٢٩٩٢	٢٩٩٩	٣٠٠٦	٣٠١٣	١	١	٢	٣	٣	٤	٥	٥	٦
٠٫٤٨	٣٠٢٠	٣٠٢٧	٣٠٣٤	٣٠٤١	٣٠٤٨	٣٠٥٥	٣٠٦٢	٣٠٦٩	٣٠٧٦	٣٠٨٣	١	١	٢	٣	٤	٤	٥	٦	٦
٠٫٤٩	٣٠٩٠	٣٠٩٧	٣١٠٥	٣١١٢	٣١١٩	٣١٢٦	٣١٣٣	٣١٤١	٣١٤٨	٣١٥٥	١	١	٢	٣	٤	٤	٥	٦	٦

(تابع) جدول الأعداد المقابلة للوغاريتمات

	٠	١	٢	٣	٤	٥	٦	٧	٨	٩	الفروق ١	٢	٣	٤	٥	٦	٧	٨	٩
٠٫٥٠	٣١٦٢	٣١٧٠	٣١٧٧	٣١٨٤	٣١٩٢	٣١٩٩	٣٢٠٦	٣٢١٤	٣٢٢١	٣٢٢٨	١	١	٢	٣	٤	٤	٥	٦	٧
٠٫٥١	٣٢٣٦	٣٢٤٣	٣٢٥١	٣٢٥٨	٣٢٦٦	٣٢٧٣	٣٢٨١	٣٢٨٩	٣٢٩٦	٣٣٠٤	١	٢	٢	٣	٤	٥	٥	٦	٧
٠٫٥٢	٣٣١١	٣٣١٩	٣٣٢٧	٣٣٣٤	٣٣٤٢	٣٣٥٠	٣٣٥٧	٣٣٦٥	٣٣٧٣	٣٣٨١	١	٢	٢	٣	٤	٥	٥	٦	٧
٠٫٥٣	٣٣٨٨	٣٣٩٦	٣٤٠٤	٣٤١٢	٣٤٢٠	٣٤٢٨	٣٤٣٦	٣٤٤٣	٣٤٥١	٣٤٥٩	١	٢	٢	٣	٤	٥	٦	٦	٧
٠٫٥٤	٣٤٦٧	٣٤٧٥	٣٤٨٣	٣٤٩١	٣٤٩٩	٣٥٠٨	٣٥١٦	٣٥٢٤	٣٥٣٢	٣٥٤٠	١	٢	٢	٣	٤	٥	٦	٦	٧
٠٫٥٥	٣٥٤٨	٣٥٥٦	٣٥٦٥	٣٥٧٣	٣٥٨١	٣٥٨٩	٣٥٩٧	٣٦٠٦	٣٦١٤	٣٦٢٢	١	٢	٢	٣	٤	٥	٦	٧	٧
٠٫٥٦	٣٦٣١	٣٦٣٩	٣٦٤٨	٣٦٥٦	٣٦٦٤	٣٦٧٣	٣٦٨١	٣٦٩٠	٣٦٩٨	٣٧٠٧	١	٢	٣	٣	٤	٥	٦	٧	٨
٠٫٥٧	٣٧١٥	٣٧٢٤	٣٧٣٣	٣٧٤١	٣٧٥٠	٣٧٥٨	٣٧٦٧	٣٧٧٦	٣٧٨٤	٣٧٩٣	١	٢	٣	٣	٤	٥	٦	٧	٨
٠٫٥٨	٣٨٠٢	٣٨١١	٣٨١٩	٣٨٢٨	٣٨٣٧	٣٨٤٦	٣٨٥٥	٣٨٦٤	٣٨٧٣	٣٨٨٢	١	٢	٣	٤	٤	٥	٦	٧	٨
٠٫٥٩	٣٨٩٠	٣٨٩٩	٣٩٠٨	٣٩١٧	٣٩٢٦	٣٩٣٦	٣٩٤٥	٣٩٥٤	٣٩٦٣	٣٩٧٢	١	٢	٣	٤	٥	٥	٦	٧	٨
٠٫٦٠	٣٩٨١	٣٩٩٠	٣٩٩٩	٤٠٠٩	٤٠١٨	٤٠٢٧	٤٠٣٦	٤٠٤٦	٤٠٥٥	٤٠٦٤	١	٢	٣	٤	٥	٦	٦	٧	٨
٠٫٦١	٤٠٧٤	٤٠٨٣	٤٠٩٣	٤١٠٢	٤١١١	٤١٢١	٤١٣٠	٤١٤٠	٤١٥٠	٤١٥٩	١	٢	٣	٤	٥	٦	٧	٨	٩
٠٫٦٢	٤١٦٩	٤١٧٨	٤١٨٨	٤١٩٨	٤٢٠٧	٤٢١٧	٤٢٢٧	٤٢٣٦	٤٢٤٦	٤٢٥٦	١	٢	٣	٤	٥	٦	٧	٨	٩
٠٫٦٣	٤٢٦٦	٤٢٧٦	٤٢٨٥	٤٢٩٥	٤٣٠٥	٤٣١٥	٤٣٢٥	٤٣٣٥	٤٣٤٥	٤٣٥٥	١	٢	٣	٤	٥	٦	٧	٨	٩
٠٫٦٤	٤٣٦٥	٤٣٧٥	٤٣٨٥	٤٣٩٥	٤٤٠٦	٤٤١٦	٤٤٢٦	٤٤٣٦	٤٤٤٦	٤٤٥٧	١	٢	٣	٤	٥	٦	٧	٨	٩
٠٫٦٥	٤٤٦٧	٤٤٧٧	٤٤٨٧	٤٤٩٨	٤٥٠٨	٤٥١٩	٤٥٢٩	٤٥٣٩	٤٥٥٠	٤٥٦٠	١	٢	٣	٤	٥	٦	٧	٨	٩
٠٫٦٦	٤٥٧١	٤٥٨١	٤٥٩٢	٤٦٠٣	٤٦١٣	٤٦٢٤	٤٦٣٤	٤٦٤٥	٤٦٥٦	٤٦٦٧	١	٢	٣	٤	٥	٦	٧	٩	١٠
٠٫٦٧	٤٦٧٧	٤٦٨٨	٤٦٩٩	٤٧١٠	٤٧٢١	٤٧٣٢	٤٧٤٢	٤٧٥٣	٤٧٦٤	٤٧٧٥	١	٢	٣	٤	٥	٧	٨	٩	١٠
٠٫٦٨	٤٧٨٦	٤٧٩٧	٤٨٠٨	٤٨١٩	٤٨٣١	٤٨٤٢	٤٨٥٣	٤٨٦٤	٤٨٧٥	٤٨٨٧	١	٢	٣	٤	٦	٧	٨	٩	١٠
٠٫٦٩	٤٨٩٨	٤٩٠٩	٤٩٢٠	٤٩٣٢	٤٩٤٣	٤٩٥٥	٤٩٦٦	٤٩٧٧	٤٩٨٩	٥٠٠٠	١	٢	٣	٥	٦	٧	٨	٩	١٠
٠٫٧٠	٥٠١٢	٥٠٢٣	٥٠٣٥	٥٠٤٧	٥٠٥٨	٥٠٧٠	٥٠٨٢	٥٠٩٣	٥١٠٥	٥١١٧	١	٢	٤	٥	٦	٧	٨	٩	١١
٠٫٧١	٥١٢٩	٥١٤٠	٥١٥٢	٥١٦٤	٥١٧٦	٥١٨٨	٥٢٠٠	٥٢١٢	٥٢٢٤	٥٢٣٦	١	٢	٤	٥	٦	٧	٨	١٠	١١
٠٫٧٢	٥٢٤٨	٥٢٦٠	٥٢٧٢	٥٢٨٤	٥٢٩٧	٥٣٠٩	٥٣٢١	٥٣٣٣	٥٣٤٦	٥٣٥٨	١	٢	٤	٥	٦	٧	٩	١٠	١١
٠٫٧٣	٥٣٧٠	٥٣٨٣	٥٣٩٥	٥٤٠٨	٥٤٢٠	٥٤٣٣	٥٤٤٥	٥٤٥٨	٥٤٧٠	٥٤٨٣	١	٣	٤	٥	٦	٨	٩	١٠	١١
٠٫٧٤	٥٤٩٥	٥٥٠٨	٥٥٢١	٥٥٣٤	٥٥٤٦	٥٥٥٩	٥٥٧٢	٥٥٨٥	٥٥٩٨	٥٦١٠	١	٣	٤	٥	٦	٨	٩	١٠	١٢
٠٫٧٥	٥٦٢٣	٥٦٣٦	٥٦٤٩	٥٦٦٢	٥٦٧٥	٥٦٨٩	٥٧٠٢	٥٧١٥	٥٧٢٨	٥٧٤١	١	٣	٤	٥	٧	٨	٩	١٠	١٢
٠٫٧٦	٥٧٥٤	٥٧٦٨	٥٧٨١	٥٧٩٤	٥٨٠٨	٥٨٢١	٥٨٣٤	٥٨٤٨	٥٨٦١	٥٨٧٥	١	٣	٤	٥	٧	٨	٩	١١	١٢
٠٫٧٧	٥٨٨٨	٥٩٠٢	٥٩١٦	٥٩٢٩	٥٩٤٣	٥٩٥٧	٥٩٧٠	٥٩٨٤	٥٩٩٨	٦٠١٢	١	٣	٤	٥	٧	٨	١٠	١١	١٢
٠٫٧٨	٦٠٢٦	٦٠٣٩	٦٠٥٣	٦٠٦٧	٦٠٨١	٦٠٩٥	٦١٠٩	٦١٢٤	٦١٣٨	٦١٥٢	١	٣	٤	٦	٧	٨	١٠	١١	١٣
٠٫٧٩	٦١٦٦	٦١٨٠	٦١٩٤	٦٢٠٩	٦٢٢٣	٦٢٣٧	٦٢٥٢	٦٢٦٦	٦٢٨١	٦٢٩٥	١	٣	٤	٦	٧	٩	١٠	١١	١٣
٠٫٨٠	٦٣١٠	٦٣٢٤	٦٣٣٩	٦٣٥٣	٦٣٦٨	٦٣٨٣	٦٣٩٧	٦٤١٢	٦٤٢٧	٦٤٤٢	١	٣	٤	٦	٧	٩	١٠	١٢	١٣
٠٫٨١	٦٤٥٧	٦٤٧١	٦٤٨٦	٦٥٠١	٦٥١٦	٦٥٣١	٦٥٤٦	٦٥٦١	٦٥٧٧	٦٥٩٢	٢	٣	٥	٦	٨	٩	١١	١٢	١٤
٠٫٨٢	٦٦٠٧	٦٦٢٢	٦٦٣٧	٦٦٥٣	٦٦٦٨	٦٦٨٣	٦٦٩٩	٦٧١٤	٦٧٣٠	٦٧٤٥	٢	٣	٥	٦	٨	٩	١١	١٢	١٤
٠٫٨٣	٦٧٦١	٦٧٧٦	٦٧٩٢	٦٨٠٨	٦٨٢٣	٦٨٣٩	٦٨٥٥	٦٨٧١	٦٨٨٧	٦٩٠٢	٢	٣	٥	٦	٨	٩	١١	١٣	١٤
٠٫٨٤	٦٩١٨	٦٩٣٤	٦٩٥٠	٦٩٦٦	٦٩٨٢	٦٩٩٨	٧٠١٥	٧٠٣١	٧٠٤٧	٧٠٦٣	٢	٣	٥	٦	٨	١٠	١١	١٣	١٥
٠٫٨٥	٧٠٧٩	٧٠٩٦	٧١١٢	٧١٢٩	٧١٤٥	٧١٦١	٧١٧٨	٧١٩٤	٧٢١١	٧٢٢٨	٢	٣	٥	٧	٨	١٠	١٢	١٣	١٥
٠٫٨٦	٧٢٤٤	٧٢٦١	٧٢٧٨	٧٢٩٥	٧٣١١	٧٣٢٨	٧٣٤٥	٧٣٦٢	٧٣٧٩	٧٣٩٦	٢	٣	٥	٧	٨	١٠	١٢	١٣	١٥
٠٫٨٧	٧٤١٣	٧٤٣٠	٧٤٤٧	٧٤٦٤	٧٤٨٢	٧٤٩٩	٧٥١٦	٧٥٣٤	٧٥٥١	٧٥٦٨	٢	٣	٥	٧	٩	١٠	١٢	١٤	١٦
٠٫٨٨	٧٥٨٦	٧٦٠٣	٧٦٢١	٧٦٣٨	٧٦٥٦	٧٦٧٤	٧٦٩١	٧٧٠٩	٧٧٢٧	٧٧٤٥	٢	٤	٥	٧	٩	١١	١٢	١٤	١٦
٠٫٨٩	٧٧٦٢	٧٧٨٠	٧٧٩٨	٧٨١٦	٧٨٣٤	٧٨٥٢	٧٨٧٠	٧٨٨٩	٧٩٠٧	٧٩٢٥	٢	٤	٥	٧	٩	١١	١٣	١٤	١٦
٠٫٩٠	٧٩٤٣	٧٩٦٢	٧٩٨٠	٧٩٩٨	٨٠١٧	٨٠٣٥	٨٠٥٤	٨٠٧٢	٨٠٩١	٨١١٠	٢	٤	٦	٧	٩	١١	١٣	١٥	١٧
٠٫٩١	٨١٢٨	٨١٤٧	٨١٦٦	٨١٨٥	٨٢٠٤	٨٢٢٢	٨٢٤١	٨٢٦٠	٨٢٧٩	٨٢٩٩	٢	٤	٦	٨	٩	١١	١٣	١٥	١٧
٠٫٩٢	٨٣١٨	٨٣٣٧	٨٣٥٦	٨٣٧٥	٨٣٩٥	٨٤١٤	٨٤٣٣	٨٤٥٣	٨٤٧٢	٨٤٩٢	٢	٤	٦	٨	١٠	١٢	١٤	١٥	١٧
٠٫٩٣	٨٥١١	٨٥٣١	٨٥٥١	٨٥٧٠	٨٥٩٠	٨٦١٠	٨٦٣٠	٨٦٥٠	٨٦٧٠	٨٦٩٠	٢	٤	٦	٨	١٠	١٢	١٤	١٦	١٨
٠٫٩٤	٨٧١٠	٨٧٣٠	٨٧٥٠	٨٧٧٠	٨٧٩٠	٨٨١٠	٨٨٣١	٨٨٥١	٨٨٧٢	٨٨٩٢	٢	٤	٦	٨	١٠	١٢	١٤	١٦	١٨
٠٫٩٥	٨٩١٣	٨٩٣٣	٨٩٥٤	٨٩٧٤	٨٩٩٥	٩٠١٦	٩٠٣٦	٩٠٥٧	٩٠٧٨	٩٠٩٩	٢	٤	٦	٨	١٠	١٢	١٥	١٧	١٩
٠٫٩٦	٩١٢٠	٩١٤١	٩١٦٢	٩١٨٣	٩٢٠٤	٩٢٢٦	٩٢٤٧	٩٢٦٨	٩٢٩٠	٩٣١١	٢	٤	٦	٨	١١	١٣	١٥	١٧	١٩
٠٫٩٧	٩٣٣٣	٩٣٥٤	٩٣٧٦	٩٣٩٧	٩٤١٩	٩٤٤١	٩٤٦٢	٩٤٨٤	٩٥٠٦	٩٥٢٨	٢	٤	٧	٩	١١	١٣	١٥	١٧	٢٠
٠٫٩٨	٩٥٥٠	٩٥٧٢	٩٥٩٤	٩٦١٦	٩٦٣٨	٩٦٦١	٩٦٨٣	٩٧٠٥	٩٧٢٧	٩٧٥٠	٢	٤	٧	٩	١١	١٣	١٦	١٨	٢٠
٠٫٩٩	٩٧٧٢	٩٧٩٥	٩٨١٧	٩٨٤٠	٩٨٦٣	٩٨٨٦	٩٩٠٨	٩٩٣١	٩٩٥٤	٩٩٧٧	٢	٥	٧	٩	١١	١٤	١٦	١٨	٢٠

جدول (م - ٨) : توزيع ستيودنت - ت

درجات الحرية	مستويات الثقة					
	أ ,٥٠ ب ,٢٥	,٤٠ ,٢٠	,٢٠ ,١٠	,١٠ ,٠٥	,٠٥ ,٠٢٥	,٠١ ,٠٠٥
١	١,٠٠	١,٣٨	٣,٠٨	٦,٣١	١٢,٧١	٦٣,٦٦
٢	,٨٢	١,٠٦	١,٨٩	٢,٩٢	٤,٣٠	٩,٩٣
٣	,٧٦	,٩٨	١,٦٤	٢,٣٥	٣,١٨	٥,٨٤
٤	,٧٤	,٩٤	١,٥٣	٢,١٣	٢,٧٨	٤,٦٠
٥	,٧٣	,٩٢	١,٤٨	٢,٠٢	٢,٥٧	٤,٠٣
٦	,٧٢	,٩١	١,٤٤	١,٩٤	٢,٤٥	٣,٧١
٧	,٧١	,٩٠	١,٤٢	١,٨٩	٢,٣٦	٣,٥٠
٨	,٧١	,٨٩	١,٤٠	١,٨٦	٢,٣١	٣,٣٦
٩	,٧٠	,٨٨	١,٣٨	١,٨٣	٢,٢٦	٣,٢٥
١٠	,٧٠	,٨٨	١,٣٧	١,٨١	٢,٢٣	٣,١٧
١١	,٧٠	,٨٨	١,٣٦	١,٨٠	٢,٢٠	٣,١١
١٢	,٦٩	,٨٧	١,٣٦	١,٧٨	٢,١٨	٣,٠٦
١٣	,٦٩	,٨٧	١,٣٥	١,٧٧	٢,١٦	٣,٠١
١٤	,٦٩	,٨٧	١,٣٤	١,٧٦	٢,١٤	٢٠٩٨
١٥	,٦٩	,٨٧	١,٣٤	١,٧٥	٢,١٣	٢,٩٥
١٦	,٦٩	,٨٧	١,٣٤	١,٧٥	٢,١٢	٢,٩٢
١٧	,٦٩	,٨٦	١,٣٣	١,٧٤	٢,١١	٢,٩٠
١٨	,٦٩	,٨٦	١,٣٣	١,٧٣	٢,١٠	٢,٨٨
١٩	,٦٩	,٨٦	١,٣٣	١,٧٣	٢,٠٩	٢,٨٦
٢٠	,٦٩	,٨٦	١,٣٢	١,٧٢	٢,٠٩	٢,٨٤
٢٢	,٦٩	,٨٦	١,٣٢	١,٧٢	٢,٠٧	٢,٨٢
٢٤	,٦٨	,٨٦	١,٣٢	١,٧١	٢,٠٦	٢,٨٠
٢٦	,٦٨	,٨٦	١,٣١	١,٧١	٢,٠٦	٢,٧٨
٢٨	,٦٨	,٨٥	١,٣١	١,٧٠	٢,٠٥	٢,٧٦
٣٠	,٦٨	,٨٥	١,٣١	١,٧٠	٢,٠٤	٢,٧٥
٣٥	,٦٨	,٨٥	١,٣١	١,٦٩	٢,٠٣	٢,٧٢
٤٠	,٦٨	,٨٥	١,٣٠	١,٦٨	٢,٠٢	٢,٧٠
٦٠	,٦٨	,٨٥	١,٣٠	١,٦٧	٢,٠٠	٢,٦٦
٨٠	,٦٨	,٨٥	١,٢٩	١,٦٧	١,٩٩	٢,٦٤
١٠٠	,٦٨	,٨٥	١,٢٩	١,٦٦	١,٩٨	٢,٦٣
∞	,٦٧	,٨٤	١,٢٨	١,٦٤	١,٩٦	٢,٥٧

أ - اختبار ثنائي الطرف

ب - اختبار أحادي الطرف

جدول رقم (م ٩) : جدول ف بمستوى دلالة ٠٥و (العادى) ، ٠١و (السفلى)

درجات الحرية للتباين الأصغر	درجات الحرية للتباين الأكبر ١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	٢٠	٣٠	٤٠	٥٠	١٠٠	∞
١	١٦١ ٤,٠٥٢	٢٠٠ ٤,٩٩٩	٢١٦ ٥,٤٠٣	٢٢٥ ٥,٦٢٥	٢٣٠ ٥,٧٦٤	٢٣٤ ٥,٨٥٩	٢٣٧ ٥,٩٢٨	٢٣٩ ٥,٩٨١	٢٤١ ٦,٠٢٢	٢٤٢ ٦,٠٥٦	٢٤٨ ٦,٢٠٨	٢٥٠ ٦,٢٥٨	٢٥١ ٦,٢٨٦	٢٥٢ ٦,٣٠٢	٢٥٣ ٦,٣٣٤	٢٥٤ ٦,٣٦٦
٢	١٨,٥١ ٩٨,٤٩	١٩,٠٠ ٩٩,٠٠	١٩,١٦ ٩٩,١٧	١٩,٢٥ ٩٩,٢٥	١٩,٣٠ ٩٩,٣٠	١٩,٣٣ ٩٩,٣٣	١٩,٣٦ ٩٩,٣٤	١٩,٣٧ ٩٩,٣٦	١٩,٣٨ ٩٩,٣٨	١٩,٣٩ ٩٩,٤٠	١٩,٤٤ ٩٩,٤٥	١٩,٤٦ ٩٩,٤٧	١٩,٤٧ ٩٩,٤٨	١٩,٤٧ ٩٩,٤٨	١٩,٤٩ ٩٩,٤٩	١٩,٥٠ ٩٩,٥٠
٣	١٠,١٣ ٣٤,١٢	٩,٥٥ ٣٠,٨٢	٩,٢٨ ٢٩,٤٦	٩,١٢ ٢٨,٧١	٩,٠١ ٢٨,٢٤	٨,٩٤ ٢٧,٩١	٨,٨٨ ٢٧,٦٧	٨,٨٤ ٢٧,٤٩	٨,٨١ ٢٧,٣٤	٨,٧٨ ٢٧,٢٣	٨,٦٦ ٢٦,٦٩	٨,٦٢ ٢٦,٥٠	٨,٦٠ ٢٦,٤١	٨,٥٨ ٢٦,٣٥	٨,٥٦ ٢٦,٢٣	٨,٥٣ ٢٦,١٢
٤	٧,٧١ ٢١,٢٠	٦,٩٤ ١٨,٠٠	٦,٥٩ ١٦,٦٩	٦,٣٩ ١٥,٩٨	٦,٢٦ ١٥,٥٢	٦,١٦ ١٥,٢١	٦,٠٩ ١٤,٩٨	٦,٠٤ ١٤,٨٠	٦,٠٠ ١٤,٦٦	٥,٩٦ ١٤,٥٤	٥,٨٠ ١٤,٠٢	٥,٧٤ ١٣,٨٣	٥,٧١ ١٣,٧٤	٥,٧٠ ١٣,٦٩	٥,٦٦ ١٣,٥٧	٥,٦٣ ١٣,٤٦
٥	٦,٦١ ١٦,٢٦	٥,٧٩ ١٣,٢٧	٥,٤١ ١٢,٠٦	٥,١٩ ١١,٣٩	٥,٠٥ ١٠,٩٧	٤,٩٥ ١٠,٦٧	٤,٨٨ ١٠,٤٥	٤,٨٢ ١٠,٢٧	٤,٧٨ ١٠,١٥	٤,٧٤ ١٠,٠٥	٤,٥٦ ٩,٥٥	٤,٥٠ ٩,٣٨	٤,٤٦ ٩,٢٩	٤,٤٤ ٩,٢٤	٤,٤٠ ٩,١٣	٤,٣٦ ٩,٠٢
٦	٥,٩٩ ١٣,٧٤	٥,١٤ ١٠,٩٢	٤,٧٦ ٩,٧٨	٤,٥٣ ٩,١٥	٤,٣٩ ٨,٧٥	٤,٢٨ ٨,٤٧	٤,٢١ ٧,٢٦	٤,١٥ ٨,١٠	٤,١٠ ٧,٩٨	٤,٠٦ ٧,٨٧	٣,٨٧ ٧,٣٩	٣,٨١ ٧,٢٣	٣,٧٧ ٧,١٤	٣,٧٥ ٧,٠٩	٣,٧١ ٦,٩٩	٣,٦٧ ٦,٨٨
٧	٥,٥٩ ١٢,٢٥	٤,٧٤ ٩,٥٥	٤,٣٥ ٨,٤٥	٤,١٢ ٧,٨٥	٣,٩٧ ٧,٤٦	٣,٨٧ ٧,١٩	٣,٧٩ ٧,٠٠	٣,٧٣ ٦,٨٤	٣,٦٨ ٦,٧١	٣,٦٣ ٦,٦٢	٣,٤٤ ٦,١٥	٣,٣٨ ٥,٩٨	٣,٣٤ ٥,٩٠	٣,٣٢ ٥,٨٥	٣,٢٨ ٥,٧٥	٣,٢٣ ٥,٦٥
٨	٥,٣٢ ١١,٢٦	٤,٤٦ ٨,٦٥	٤,٠٧ ٧,٥٩	٣,٨٤ ٧,٠١	٣,٦٩ ٦,٦٣	٣,٥٨ ٦,٣٧	٣,٥٠ ٦,١٩	٣,٤٤ ٦,٠٣	٣,٣٩ ٥,٩١	٣,٣٤ ٥,٨٢	٣,١٥ ٥,٣٦	٣,٠٨ ٥,٢٠	٣,٠٥ ٥,١١	٣,٠٣ ٥,٠٦	٢,٩٨ ٤,٩٦	٢,٩٣ ٤,٨٦

(تابع) جدول رقم (٣-٩)

درجات الحرية للتباين الأصغر	درجات الحرية للتباين الأكبر: ١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	٢٠	٣٠	٤٠	٥٠	١٠٠	∞
٩	٥٫١٢	٤٫٢٦	٣٫٨٦	٣٫٦٣	٣٫٤٨	٣٫٣٧	٣٫٢٩	٣٫٢٣	٣٫١٨	٣٫١٣	٢٫٩٣	٢٫٨٦	٢٫٨٢	٢٫٨٠	٢٫٧٦	٢٫٧١
	١٠٫٥٦	٨٫٠٢	٦٫٩٩	٦٫٤٢	٦٫٠٦	٥٫٨٠	٥٫٦٢	٥٫٤٧	٥٫٣٥	٥٫٢٦	٤٫٨٠	٤٫٦٤	٤٫٥٦	٤٫٥١	٤٫٤١	٤٫٣١
١٠	٤٫٩٦	٤٫١٠	٣٫٧١	٣٫٤٨	٣٫٣٣	٣٫٢٢	٣٫١٤	٣٫٠٧	٣٫٠٢	٢٫٩٧	٢٫٧٧	٢٫٧٠	٢٫٦٧	٢٫٦٤	٢٫٥٩	٢٫٥٤
	١٠٫٠٤	٧٫٥٦	٦٫٥٥	٥٫٩٩	٥٫٦٤	٥٫٣٩	٥٫٢١	٥٫٠٦	٤٫٩٥	٤٫٨٥	٤٫٤١	٤٫٢٥	٤٫١٧	٤٫١٢	٤٫٠١	٣٫٩١
١١	٤٫٨٤	٣٫٩٨	٣٫٥٩	٣٫٣٦	٣٫٢٠	٣٫٠٩	٣٫٠١	٢٫٩٥	٢٫٩٠	٢٫٨٦	٢٫٦٥	٢٫٥٧	٢٫٥٣	٢٫٥٠	٢٫٤٥	٢٫٤٠
	٩٫٦٥	٧٫٢٠	٦٫٢٢	٥٫٦٧	٥٫٣٢	٥٫٠٧	٤٫٨٨	٤٫٧٤	٤٫٦٣	٤٫٥٤	٤٫١٠	٣٫٩٤	٣٫٨٦	٣٫٨٠	٣٫٧٠	٣٫٦٠
١٢	٤٫٧٥	٣٫٨٨	٣٫٤٩	٣٫٢٦	٣٫١١	٣٫٠٠	٢٫٩٢	٢٫٨٥	٢٫٨٠	٢٫٧٦	٢٫٥٤	٢٫٤٦	٢٫٤٢	٢٫٤٠	٢٫٣٥	٢٫٣٠
	٩٫٣٣	٦٫٩٣	٥٫٩٥	٥٫٤١	٥٫٠٦	٤٫٨٢	٤٫٦٥	٤٫٥٠	٤٫٣٩	٤٫٣٠	٣٫٨٦	٣٫٧٠	٣٫٦١	٣٫٥٦	٣٫٤٦	٣٫٣٦
١٣	٤٫٦٧	٣٫٨٠	٣٫٤١	٣٫١٨	٣٫٠٢	٢٫٩٢	٢٫٨٤	٢٫٧٧	٢٫٧٢	٢٫٦٧	٢٫٤٦	٢٫٣٨	٢٫٣٤	٢٫٣٢	٢٫٢٦	٢٫٢١
	٩٫٠٧	٦٫٧٠	٥٫٧٤	٥٫٢٠	٤٫٨٦	٤٫٦٢	٤٫٤٤	٤٫٣٠	٤٫١٩	٤٫١٠	٣٫٦٧	٣٫٥١	٣٫٤٢	٣٫٣٧	٣٫٢٧	٣٫١٦
١٤	٤٫٦٠	٣٫٧٤	٣٫٣٤	٣٫١١	٢٫٩٦	٢٫٨٥	٢٫٧٧	٢٫٧٠	٢٫٦٥	٢٫٦٠	٢٫٣٩	٢٫٣١	٢٫٢٧	٢٫٢٤	٢٫١٩	٢٫١٣
	٨٫٨٦	٦٫٥١	٥٫٥٦	٥٫٠٣	٤٫٦٩	٤٫٤٦	٤٫٢٨	٤٫١٤	٤٫٠٣	٣٫٩٤	٣٫٥١	٣٫٣٤	٣٫٢٦	٣٫٢١	٣٫١١	٣٫٠٠
١٥	٤٫٥٤	٣٫٦٨	٣٫٢٩	٣٫٠٦	٢٫٩٠	٢٫٧٩	٢٫٧٠	٢٫٦٤	٢٫٥٩	٢٫٥٥	٢٫٣٣	٢٫٢٥	٢٫٢١	٢٫١٨	٢٫١٢	٢٫٠٧
	٨٫٦٨	٦٫٣٦	٥٫٤٢	٤٫٨٩	٤٫٥٦	٤٫٣٢	٤٫١٤	٤٫٠٠	٣٫٨٩	٣٫٨٠	٣٫٣٦	٣٫٢٠	٣٫١٢	٣٫٠٧	٢٫٩٧	٢٫٨٧
١٦	٤٫٤٩	٣٫٦٣	٣٫٢٤	٣٫٠١	٢٫٨٥	٢٫٧٤	٢٫٦٦	٢٫٥٩	٢٫٥٤	٢٫٤٩	٢٫٢٨	٢٫٢٠	٢٫١٦	٢٫١٣	٢٫٠٧	٢٫٠١
	٨٫٥٣	٦٫٢٣	٥٫٢٩	٤٫٧٧	٤٫٤٤	٤٫٢٠	٤٫٠٣	٣٫٨٩	٣٫٧٨	٣٫٦٩	٣٫٢٥	٣٫١٠	٣٫٠١	٢٫٩٦	٢٫٨٦	٢٫٧٥

تابع جدول (م-٩)

درجات الحرية للتباين الأصغر	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	٢٠	٣٠	٤٠	٥٠	١٠٠	∞
	درجات الحرية للتباين الأكبر															
١٧	٤,٤٥ ٨,٤٠	٣,٥٩ ٦,١١	٣,٢٠ ٥,١٨	٢,٩٦ ٤,٦٧	٢,٨١ ٤,٣٤	٢,٧٠ ٤,١٠	٢,٦٢ ٣,٩٣	٢,٥٥ ٣,٧٩	٢,٥٠ ٣,٦٨	٢,٤٥ ٣,٥٩	٢,٢٣ ٣,١٦	٢,١٥ ٣,٠٠	٢,١١ ٢,٩٢	٢,٠٨ ٢,٨٦	٢,٠٢ ٢,٧٦	١,٩٦ ٢,٦٥
١٨	٤,٤١ ٨,٢٨	٣,٥٥ ٦,٠١	٣,١٦ ٥,٠٩	٢,٩٣ ٤,٥٨	٢,٧٧ ٤,٢٥	٢,٦٦ ٤,٠١	٢,٥٨ ٣,٨٥	٢,٥١ ٣,٧١	٢,٤٦ ٣,٦٠	٢,٤١ ٣,٥١	٢,١٩ ٣,٠٧	٢,١١ ٢,٩١	٢,٠٧ ٢,٨٣	٢,٠٤ ٢,٧٨	١,٩٨ ٢,٦٨	١,٩٢ ٢,٥٧
١٩	٤,٣٨ ٨,١٨	٣,٥٢ ٥,٩٣	٣,١٣ ٥,٠١	٢,٩٠ ٤,٥٠	٢,٧٤ ٤,١٧	٢,٦٣ ٣,٩٤	٢,٥٥ ٣,٧٧	٢,٤٨ ٣,٦٣	٢,٤٣ ٣,٥٢	٢,٣٨ ٣,٤٣	٢,١٥ ٣,٠٠	٢,٠٧ ٢,٨٤	٢,٠٢ ٢,٧٦	٢,٠٠ ٢,٧٠	١,٩٤ ٢,٦٠	١,٨٨ ٢,٤٩
٢٠	٤,٣٥ ٨,١٠	٣,٤٩ ٥,٨٥	٣,١٠ ٤,٩٤	٢,٨٧ ٤,٤٣	٢,٧١ ٤,١٠	٢,٦٠ ٣,٨٧	٢,٥٢ ٣,٧١	٢,٤٥ ٣,٥٦	٢,٤٠ ٣,٤٥	٢,٣٥ ٣,٣٧	٢,١٢ ٢,٩٤	٢,٠٤ ٢,٧٧	١,٩٩ ٢,٦٩	١,٩٦ ٢,٦٣	١,٩٠ ٢,٥٣	١,٨٤ ٢,٤٢
٢١	٤,٣٢ ٨,٠٢	٣,٤٧ ٥,٧٨	٣,٠٧ ٤,٨٧	٢,٨٤ ٤,٣٧	٢,٦٨ ٤,٠٤	٢,٥٧ ٣,٨١	٢,٤٩ ٣,٦٥	٢,٤٢ ٣,٥١	٢,٣٧ ٣,٤٠	٢,٣٢ ٣,٣١	٢,٠٩ ٢,٨٨	٢,٠٠ ٢,٧٢	١,٩٦ ٢,٦٣	١,٩٣ ٢,٥٨	١,٨٧ ٢,٤٧	١,٨١ ٢,٣٦
٢٢	٤,٣٠ ٧,٩٤	٣,٤٤ ٥,٧٢	٣,٠٥ ٤,٨٢	٢,٨٢ ٤,٣١	٢,٦٦ ٣,٩٩	٢,٥٥ ٣,٧٦	٢,٤٧ ٣,٥٩	٢,٤٠ ٣,٤٥	٢,٣٥ ٣,٣٥	٢,٣٠ ٣,٢٦	٢,٠٧ ٢,٨٣	١,٩٨ ٢,٦٧	١,٩٣ ٢,٥٨	١,٩١ ٢,٥٣	١,٨٤ ٢,٤٢	١,٧٨ ٢,٣١
٢٣	٤,٢٨ ٧,٨٨	٣,٤٢ ٥,٦٦	٣,٠٣ ٤,٧٦	٢,٨٠ ٤,٢٦	٢,٦٤ ٣,٩٤	٢,٥٣ ٣,٧١	٢,٤٥ ٣,٥٤	٢,٣٨ ٣,٤١	٢,٣٢ ٣,٣٠	٢,٢٨ ٣,٢١	٢,٠٤ ٢,٧٨	١,٩٦ ٢,٦٢	١,٩١ ٢,٥٣	١,٨٨ ٢,٤٨	١,٨٢ ٢,٣٧	١,٧٦ ٢,٢٦
٢٤	٤,٢٦ ٧,٨٢	٣,٤٠ ٥,٦١	٣,٠١ ٤,٧٢	٢,٧٨ ٤,٢٢	٢,٦٢ ٣,٩٠	٢,٥١ ٣,٦٧	٢,٤٣ ٣,٥٠	٢,٣٦ ٣,٣٦	٢,٣٠ ٣,٢٥	٢,٢٦ ٣,١٧	٢,٠٢ ٢,٧٤	١,٩٤ ٢,٥٨	١,٨٩ ٢,٤٩	١,٨٦ ٢,٤٤	١,٨٠ ٢,٣٣	١,٧٣ ٢,٢١

(تابع) جدول رقم (٣-٩)

درجات الحرية للتباين الاصغر	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	٢٠	٣٠	٤٠	٥٠	١٠٠	∞
	درجات الحرية للتباين الاكبر															
٢٥	٤,٢٤ ٧,٧٧	٣,٣٨ ٥,٥٧	٢,٩٩ ٤,٦٨	٢,٧٦ ٤,١٨	٢,٦٠ ٣,٨٦	٢,٤٩ ٣,٦٣	٢,٤١ ٣,٤٦	٢,٣٤ ٣,٣٢	٢,٢٨ ٣,٢١	٢,٢٤ ٣,١٣	٢,٠٠ ٢,٧٠	١,٩٢ ٢,٥٤	١,٨٧ ٢,٤٥	١,٨٤ ٢,٤٠	١,٧٧ ٢,٢٩	١,٧١ ٢,١٧
٢٦	٤,٢٢ ٧,٧٢	٣,٣٧ ٥,٥٣	٢,٩٨ ٤,٦٤	٢,٧٤ ٤,١٤	٢,٥٩ ٣,٨٢	٢,٤٧ ٣,٥٩	٢,٣٩ ٣,٤٢	٢,٣٢ ٣,٢٩	٢,٢٧ ٣,١٧	٢,٢٢ ٣,٠٩	١,٩٩ ٢,٦٦	١,٩٠ ٢,٥٠	١,٨٥ ٢,٤١	١,٨٢ ٢,٣٦	١,٧٦ ٢,٢٥	١,٦٩ ٢,١٣
٢٧	٤,٢١ ٧,٦٨	٣,٣٥ ٥,٤٩	٢,٩٦ ٤,٦٠	٢,٧٣ ٤,١١	٢,٥٧ ٣,٧٩	٢,٤٦ ٣,٥٦	٢,٣٧ ٣,٣٩	٢,٣٠ ٣,٢٦	٢,٢٥ ٣,١٤	٢,٢٠ ٣,٠٦	١,٩٧ ٢,٦٣	١,٨٨ ٢,٤٧	١,٨٤ ٢,٣٨	١,٨٠ ٢,٣٣	١,٧٤ ٢,٢١	١,٦٧ ٢,١٠
٢٨	٤,٢٠ ٧,٦٤	٣,٣٤ ٥,٤٥	٢,٩٥ ٤,٥٧	٢,٧١ ٤,٠٧	٢,٥٦ ٣,٧٦	٢,٤٤ ٣,٥٣	٢,٣٦ ٣,٣٦	٢,٢٩ ٣,٢٣	٢,٢٤ ٣,١١	٢,١٩ ٣,٠٣	١,٩٦ ٢,٦٠	١,٨٧ ٢,٤٤	١,٨١ ٢,٣٥	١,٧٨ ٢,٣٠	١,٧٢ ٢,١٨	١,٦٥ ٢,٠٦
٢٩	٤,١٨ ٧,٦٠	٣,٣٣ ٥,٤٢	٢,٩٣ ٤,٥٤	٢,٧٠ ٤,٠٤	٢,٥٤ ٣,٧٣	٢,٤٣ ٣,٥٠	٢,٣٥ ٣,٣٣	٢,٢٨ ٣,٢٠	٢,٢٢ ٣,٠٨	٢,١٨ ٣,٠٠	١,٩٤ ٢,٥٧	١,٨٥ ٢,٤١	١,٨٠ ٢,٣٢	١,٧٧ ٢,٢٧	١,٧١ ٢,١٥	١,٦٤ ٢,٠٣
٣٠	٤,١٧ ٧,٥٦	٣,٣٢ ٥,٣٩	٢,٩٢ ٤,٥١	٢,٦٩ ٤,٠٢	٢,٥٣ ٣,٧٠	٢,٤٢ ٣,٤٧	٢,٣٤ ٣,٣٠	٢,٢٧ ٣,١٧	٢,٢١ ٣,٠٦	٢,١٦ ٢,٩٨	١,٩٣ ٢,٥٥	١,٨٤ ٢,٣٨	١,٧٩ ٢,٢٩	١,٧٦ ٢,٢٤	١,٦٩ ٢,١٣	١,٦٢ ٢,٠١
٣٢	٤,١٥ ٧,٥٠	٣,٣٠ ٥,٣٤	٢,٩٠ ٤,٤٦	٢,٦٧ ٣,٩٧	٢,٥١ ٣,٦٦	٢,٤٠ ٣,٤٢	٢,٣٢ ٣,٢٥	٢,٢٥ ٣,١٢	٢,١٩ ٣,٠١	٢,١٤ ٢,٩٤	١,٩١ ٢,٥١	١,٨٢ ٢,٣٤	١,٧٦ ٢,٢٥	١,٧٤ ٢,٢٠	١,٦٧ ٢,٠٨	١,٥٩ ١,٩٦
٣٤	٤,١٣ ٧,٤٤	٣,٢٨ ٥,٢٩	٢,٨٨ ٤,٤٢	٢,٦٥ ٣,٩٣	٢,٤٩ ٣,٦١	٢,٣٨ ٣,٣٨	٢,٣٠ ٣,٢١	٢,٢٣ ٣,٠٨	٢,١٧ ٢,٩٧	٢,١٢ ٢,٨٩	١,٨٩ ٢,٤٧	١,٨٠ ٢,٣٠	١,٧٤ ٢,٢١	١,٧١ ٢,١٥	١,٦٤ ٢,٠٤	١,٥٧ ١,٩١

(تابع) جدول رقم (٣ - ٩)

درجات الحرية للتباين الاصغر	درجات الحرية للتباين الاكبر ١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	٢٠	٣٠	٤٠	٥٠	١٠٠	∞
٣٦	٤٫١١ ٧٫٣٩	٣٫٢٦ ٥٫٢٥	٢٫٨٦ ٤٫٣٨	٢٫٦٣ ٣٫٨٩	٢٫٤٨ ٣٫٥٨	٢٫٣٦ ٣٫٣٥	٢٫٢٨ ٣٫١٨	٢٫٢١ ٣٫٠٤	٢٫١٥ ٢٫٩٤	٢٫١٠ ٢٫٨٦	[illegible] ٢٫٤٣	١٫٧٨ ٢٫٢٦	١٫٧٢ ٢٫١٧	١٫٦٩ ٢٫١٢	١٫٦٢ ٢٫٠٠	١٫٥٥ ١٫٨٧
٣٨	٤٫١٠ ٧٫٣٥	٣٫٢٥ ٥٫٢١	٢٫٨٥ ٤٫٣٤	٢٫٦٢ ٣٫٨٦	٢٫٤٦ ٣٫٥٤	٢٫٣٥ ٣٫٣٢	٢٫٢٦ ٣٫١٥	٢٫١٩ ٣٫٠٢	٢٫١٤ ٢٫٩١	٢٫٠٩ ٢٫٨٢	١٫٨٥ ٢٫٤٠	١٫٧٦ ٢٫٢٢	١٫٧١ ٢٫١٤	١٫٦٧ ٢٫٠٨	١٫٦٠ ١٫٩٧	١٫٥٣ ١٫٨٤
٤٠	٤٫٠٨ ٧٫٣١	٣٫٢٣ ٥٫١٨	٢٫٨٤ ٤٫٣١	٢٫٦١ ٣٫٨٣	٢٫٤٥ ٣٫٥١	٢٫٣٤ ٣٫٢٩	٢٫٢٥ ٣٫١٢	٢٫١٨ ٢٫٩٩	٢٫١٢ ٢٫٨٨	٢٫٠٧ ٢٫٨٠	١٫٨٤ ٢٫٣٧	١٫٧٤ ٢٫٢٠	١٫٦٩ ٢٫١١	١٫٦٦ ٢٫٠٥	١٫٥٩ ١٫٩٤	١٫٥١ ١٫٨١
٤٢	٤٫٠٧ ٧٫٢٧	٣٫٢٢ ٥٫١٥	٢٫٨٣ ٤٫٢٩	٢٫٥٩ ٣٫٨٠	٢٫٤٤ ٣٫٤٩	٢٫٣٢ ٣٫٢٦	٢٫٢٤ ٣٫١٠	٢٫١٧ ٢٫٩٦	٢٫١١ ٢٫٨٦	٢٫٠٦ ٢٫٧٧	١٫٨٢ ٢٫٣٥	١٫٧٣ ٢٫١٧	١٫٦٨ ٢٫٠٨	١٫٦٤ ٢٫٠٢	١٫٥٧ ١٫٩١	١٫٤٩ ١٫٧٨
٤٤	٤٫٠٦ ٧٫٢٤	٣٫٢١ ٥٫١٢	٢٫٨٢ ٤٫٢٦	٢٫٥٨ ٣٫٧٨	٢٫٤٣ ٣٫٤٦	٢٫٣١ ٣٫٢٤	٢٫٢٣ ٣٫٠٧	٢٫١٦ ٢٫٩٤	٢٫١٠ ٢٫٨٤	٢٫٠٥ ٢٫٧٥	١٫٨١ ٢٫٣٢	١٫٧٢ ٢٫١٥	١٫٦٦ ٢٫٠٦	١٫٦٣ ٢٫٠٠	١٫٥٦ ١٫٨٨	١٫٤٨ ١٫٧٥
٤٦	٤٫٠٥ ٧٫٢١	٣٫٢٠ ٥٫١٠	٢٫٨١ ٤٫٢٤	٢٫٥٧ ٣٫٧٦	٢٫٤٢ ٣٫٤٤	٢٫٣٠ ٣٫٢٢	٢٫٢٢ ٣٫٠٥	٢٫١٤ ٢٫٩٢	٢٫٠٩ ٢٫٨٢	٢٫٠٤ ٢٫٧٣	١٫٨٠ ٢٫٣٠	١٫٧١ ٢٫١٣	١٫٦٥ ٢٫٠٤	١٫٦٢ ١٫٩٨	١٫٥٤ ١٫٨٦	١٫٤٦ ١٫٧٢
٤٨	٤٫٠٤ ٧٫١٩	٣٫١٩ ٥٫٠٨	٢٫٨٠ ٤٫٢٢	٢٫٥٦ ٣٫٧٤	٢٫٤١ ٣٫٤٢	٢٫٣٠ ٣٫٢٠	٢٫٢١ ٣٫٠٤	٢٫١٤ ٢٫٩٠	٢٫٠٨ ٢٫٨٠	٢٫٠٣ ٢٫٧١	١٫٧٩ ٢٫٢٨	١٫٧٠ ٢٫١١	١٫٦٤ ٢٫٠٢	١٫٦١ ١٫٩٦	١٫٥٣ ١٫٨٤	١٫٤٥ ١٫٧٠
٥٠	٤٫٠٣ ٧٫١٧	٣٫١٨ ٥٫٠٦	٢٫٧٩ ٤٫٢٠	٢٫٥٦ ٣٫٧٢	٢٫٤٠ ٣٫٤١	٢٫٢٩ ٣٫١٨	٢٫٢٠ ٣٫٠٢	٢٫١٣ ٢٫٨٨	٢٫٠٧ ٢٫٧٨	٢٫٠٢ ٢٫٧٠	١٫٧٨ ٢٫٢٦	١٫٦٩ ٢٫١٠	١٫٦٣ ٢٫٠٠	١٫٦٠ ١٫٩٤	١٫٥٢ ١٫٨٢	١٫٤٤ ١٫٦٨

(تابع) جدول رقم (٣ - ٩)

درجات الحرية للمقام	درجات الحرية للتباين الأكبر ١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	٢٠	٣٠	٤٠	٥٠	١٠٠	∞
٥٥	٤،٠٢ ٧،١٢	٣،١٧ ٥،٠١	٢،٧٨ ٤،١٦	٢،٥٤ ٣،٦٨	٢،٣٨ ٣،٣٧	٢،٢٧ ٣،١٥	٢،١٨ ٢،٩٨	٢،١١ ٢،٨٥	٢،٠٥ ٢،٧٥	٢،٠٠ ٢،٦٦	١،٧٦ ٢،٢٣	١،٦٧ ٢،٠٦	١،٦١ ١،٩٦	١،٥٨ ١،٩٠	١،٥ ١،٧٨	١،٤١ ١،٦٤
٦٠	٤،٠٠ ٧،٠٨	٣،١٥ ٤،٩٨	٢،٧٦ ٤،١٣	٢،٥٢ ٣،٦٥	٢،٣٧ ٣،٣٤	٢،٢٥ ٣،١٢	٢،١٧ ٢،٩٥	٢،١٠ ٢،٨٢	٢،٠٤ ٢،٧٢	١،٩٩ ٢،٦٣	١،٧٥ ٢،٢٠	١،٦٥ ٢،٠٣	١،٥٩ ١،٩٣	١،٥٦ ١،٨٧	١،٤٨ ١،٧٤	١،٣٩ ١،٦٠
٦٥	٣،٩٩ ٧،٠٤	٣،١٤ ٤،٩٥	٢،٧٥ ٤،١٠	٢،٥١ ٣،٦٢	٢،٣٦ ٣،٣١	٢،٢٤ ٣،٠٩	٢،١٥ ٢،٩٣	٢،٠٨ ٢،٧٩	٢،٠٢ ٢،٧٠	١،٩٨ ٢،٦١	١،٧٣ ٢،١٨	١،٦٣ ٢،٠٠	١،٥٧ ١،٩٠	١،٥٤ ١،٨٤	١،٤٦ ١،٧١	١،٣٧ ١،٥٦
٧٠	٣،٩٨ ٧،٠١	٣،١٣ ٤،٩٢	٢،٧٤ ٤،٠٨	٢،٥٠ ٣،٦٠	٢،٣٥ ٣،٢٩	٢،٢٣ ٣،٠٧	٢،١٤ ٢،٩١	٢،٠٧ ٢،٧٧	٢،٠١ ٢،٦٧	١،٩٧ ٢،٥٩	١،٧٢ ٢،١٥	١،٦٢ ١،٩٨	١،٥٦ ١،٨٨	١،٥٣ ١،٨٢	١،٤٥ ١،٦٩	١،٣٥ ١،٥٣
٨٠	٣،٩٦ ٦،٩٦	٣،١١ ٤،٨٨	٢،٧٢ ٤،٠٤	٢،٤٨ ٣،٥٦	٢،٣٣ ٣،٢٥	٢،٢١ ٣،٠٤	٢،١٢ ٢،٨٧	٢،٠٥ ٢،٧٤	١،٩٩ ٢،٦٤	١،٩٥ ٢،٥٥	١،٧٠ ٢،١١	١،٦٠ ١،٩٤	١،٥٤ ١،٨٤	١،٥١ ١،٧٨	١،٤٢ ١،٦٥	١،٣٢ ١،٤٩
١٠٠	٣،٩٤ ٦،٩٠	٣،٠٩ ٤،٨٢	٢،٧٠ ٣،٩٨	٢،٤٦ ٣،٥١	٢،٣٠ ٣،٢٠	٢،١٩ ٢،٩٩	٢،١٠ ٢،٨٢	٢،٠٣ ٢،٦٩	١،٩٧ ٢،٥٩	١،٩٢ ٢،٥١	١،٦٨ ٢،٠٦	١،٥٧ ١،٨٩	١،٥١ ١،٧٩	١،٤٨ ١،٧٣	١،٣٩ ١،٥٩	١،٢٨ ١،٤٣
١٢٥	٣،٩٢ ٦،٨٤	٣،٠٧ ٤،٧٨	٢،٦٨ ٣،٩٤	٢،٤٤ ٣،٤٧	٢،٢٩ ٣،١٧	٢،١٧ ٢،٩٥	٢،٠٨ ٢،٧٩	٢،٠١ ٢،٦٥	١،٩٥ ٢،٥٦	١،٩٠ ٢،٤٧	١،٦٥ ٢،٠٣	١،٥٥ ١،٨٥	١،٤٩ ١،٧٥	١،٤٥ ١،٦٨	١،٣٦ ١،٥٤	١،٢٥ ١،٣٧
١٥٠	٣،٩١ ٦،٨١	٣،٠٦ ٤،٧٥	٢،٦٧ ٣،٩١	٢،٤٣ ٣،٤٤	٢،٢٧ ٣،١٤	٢،١٦ ٢،٩٢	٢،٠٧ ٢،٧٦	٢،٠٠ ٢،٦٢	١،٩٤ ٢،٥٣	١،٨٩ ٢،٤٤	١،٦٤ ٢،٠٠	١،٥٤ ١،٨٣	١،٤٧ ١،٧٢	١،٤٤ ١،٦٦	١،٣٤ ١،٥١	١،٢٢ ١،٣٣

(تابع) جدول رقم (٣-٩)

درجات الحرية للتباين الأصغر	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	٢٠	٣٠	٤٠	٥٠	١٠٠	∞
	درجات الحرية للتباين الأكبر															
٢٠٠	٣,٨٩	٣,٠٤	٢,٦٥	٢,٤١	٢,٢٦	٢,١٤	٢,٠٥	١,٩٨	١,٩٢	١,٨٧	١,٦٢	١,٥٢	١,٤٥	١,٤٢	١,٣٢	١,١٩
	٦,٧٦	٤,٧١	٣,٨٨	٣,٤١	٣,١١	٢,٩٠	٢,٧٣	٢,٦٠	٢,٥٠	٢,٤١	١,٩	١,٧٩	١,٦٩	١,٦٢	١,٤٨	١,٢٨
٤٠٠	٣,٨٦	٣,٠٢	٢,٦٢	٢,٣٩	٢,٢٣	٢,١٢	٢,٠٣	١,٩٦	١,٩٠	١,٨٥	١,٦٠	١,٤٩	١,٤٢	١,٣٨	١,٢٨	١,١٣
	٦,٧٠	٤,٦٦	٣,٨٣	٣,٣٦	٣,٠٦	٢,٨٥	٢,٦٩	٢,٥٥	٢,٤٦	٢,٣٧	١,٩٢	١,٧٤	١,٦٤	١,٥٧	١,٤٢	١,١٩
١٠٠٠	٣,٨٥	٣,٠٠	٢,٦١	٢,٣٨	٢,٢٢	٢,١٠	٢,٠٢	١,٩٥	١,٨٩	١,٨٤	١,٥٨	١,٤٧	١,٤١	١,٣٦	١,٢٦	١,٠٨
	٦,٦٦	٤,٦٢	٣,٨٠	٣,٣٤	٣,٠٤	٢,٨٢	٢,٦٦	٢,٥٣	٢,٤٣	٢,٣٤	١,٨٩	١,٧١	١,٦١	١,٥٤	١,٣٨	١,١١
∞	٣,٨٤	٢,٩٩	٢,٦٠	٢,٣٧	٢,٢١	٢,٠٩	٢,٠١	١,٩٤	١,٨٨	١,٨٣	١,٥٧	١,٤٦	١,٤٠	١,٣٥	١,٢٤	١,٠٠
	٦,٦٤	٤,٦٠	٣,٧٨	٣,٣٢	٣,٠٢	٢,٨٠	٢,٦٤	٢,٥١	٢,٤١	٢,٣٢	١,٨٧	١,٦٩	١,٥٩	١,٥٢	١,٣٦	١,٠٠

جدول رقم (م - ١٠): جدول مربع كاى (χ^2)

درجات الحرية	احتمال الحصول على قيمة أعلى من كا٢									
	٠,٩٩٠	٠,٩٧٥	٠,٩٠٠	٠,٧٥٠	٠,٥٠٠	٠,٢٥٠	٠,١٠٠	٠,٠٥	٠,٠٢٥	٠,٠٠١
١	٠,٠٠٠١	٠,٠٠٤١٠	٠,٠١٥٨	٠,١٠٢	٠,٤٥٥	١,٣٢	٢,٧١	٣,٨٤	٥,٠٢	٦,٦٣
٢	٠,٠٢٠١	٠,٠٥٠٦	٠,٢١١	٠,٥٧٥	١,٣٩	٢,٧٧	٤,٦١	٥,٩٩	٧,٣٨	٩,٢١
٣	٠,١١٥	٠,٢١٦	٠,٥٨٤	١,٢١	٢,٣٧	٤,١١	٦,٢٥	٧,٨١	٩,٣٥	١١,٣
٤	٠,٢٩٧	٠,٤٨٤	١,٠٦	١,٩٢	٣,٣٦	٥,٣٩	٧,٧٨	٩,٤٩	١١,١	١٣,٣
٥	٠,٥٥٤	٠,٨٣١	١,٦١	٢,٦٧	٤,٣٥	٦,٦٣	٩,٢٤	١١,١	١٢,٨	١٥,١
٦	٠,٨٧٢	١,٢٤	٢,٢٠	٣,٤٥	٥,٣٥	٧,٨٤	١٠,٦	١٢,٦	١٤,٤	١٦,٨
٧	١,٢٤	١,٦٩	٢,٨٣	٤,٢٥	٦,٣٥	٩,٠٤	١٢,٠	١٤,١	١٦,٠	١٨,٥
٨	١,٦٥	٢,١٨	٣,٤٩	٥,٠٧	٧,٣٤	١٠,٢	١٣,٤	١٥,٥	١٧,٥	٢٠,١
٩	٢,٠٩	٢,٧٠	٤,١٧	٥,٩٠	٨,٣٤	١١,٤	١٤,٧	١٦,٩	١٩,٠	٢١,٦
١٠	٢,٥٦	٣,٢٥	٤,٨٧	٦,٧٤	٩,٣٤	١٢,٥	١٦,٠	١٨,٣	٢٠,٥	٢٣,٢
١١	٣,٠٥	٣,٨٢	٥,٥٨	٧,٥٨	١٠,٣	١٣,٧	١٧,٣	١٩,٧	٢١,٩	٢٤,٧
١٢	٣,٥٧	٤,٤٠	٦,٣٠	٨,٤٤	١١,٣	١٤,٨	١٨,٥	٢١,٠	٢٣,٣	٢٦,٢
١٣	٤,١١	٥,٠١	٧,٠٤	٩,٣٠	١٢,٣	١٦,٠	١٩,٨	٢٢,٤	٢٤,٧	٢٧,٧
١٤	٤,٦٦	٥,٦٣	٧,٧٩	١٠,٢	١٣,٣	١٧,١	٢١,١	٢٣,٧	٢٦,١	٢٩,١
١٥	٥,٢٣	٦,٢٦	٨,٥٥	١١,٠	١٤,٣	١٨,٢	٢٢,٣	٢٥,٠	٢٧,٥	٣٠,٦
١٦	٥,٨١	٦,٩١	٩,٣١	١١,٩	١٥,٣	١٩,٤	٢٣,٥	٢٦,٣	٢٨,٨	٣٢,٠
١٧	٦,٤١	٧,٥٦	١٠,١	١٢,٨	١٦,٣	٢٠,٥	٢٤,٨	٢٧,٦	٣٠,٢	٣٣,٤
١٨	٧,٠١	٨,٢٣	١٠,٩	١٣,٧	١٧,٣	٢١,٦	٢٦,٠	٢٨,٩	٣١,٥	٣٤,٨
١٩	٧,٦٣	٨,٩١	١١,٧	١٤,٦	١٨,٣	٢٢,٧	٢٧,٢	٣٠,١	٣٢,٩	٣٦,٢
٢٠	٨,٢٦	٩,٥٩	١٢,٤	١٥,٥	١٩,٣	٢٣,٨	٢٨,٤	٣١,٤	٣٤,٢	٣٧,٦
٢١	٨,٩٠	١٠,٣	١٣,٢	١٦,٣	٢٠,٣	٢٤,٩	٢٩,٦	٣٢,٧	٣٥,٥	٣٨,٩
٢٢	٩,٥٤	١١,٠	١٤,٠	١٧,٢	٢١,٣	٢٦,٠	٣٠,٨	٣٣,٩	٣٦,٨	٤٠,٣
٢٣	١٠,٢	١١,٧	١٤,٨	١٨,١	٢٢,٣	٢٧,١	٣٢,٠	٣٥,٢	٣٨,١	٤١,٦
٢٤	١٠,٩	١٢,٤	١٥,٧	١٩,٠	٢٣,٣	٢٨,٢	٣٣,٢	٣٦,٤	٣٩,٤	٤٣,٠
٢٥	١١,٥	١٣,١	١٦,٥	١٩,٩	٢٤,٣	٢٩,٣	٣٤,٤	٣٧,٧	٤٠,٦	٤٤,٣
٢٦	١٢,٢	١٣,٨	١٧,٣	٢٠,٨	٢٥,٣	٣٠,٤	٣٥,٦	٣٨,٩	٤١,٩	٤٥,٦
٢٧	١٢,٩	١٤,٦	١٨,١	٢١,٧	٢٦,٣	٣١,٥	٣٦,٧	٤٠,١	٤٣,٢	٤٧,٠
٢٨	١٣,٦	١٥,٣	١٨,٩	٢٢,٧	٢٧,٣	٣٢,٦	٣٧,٩	٤١,٣	٤٤,٥	٤٨,٣
٢٩	١٤,٣	١٦,٠	١٩,٨	٢٣,٦	٢٨,٣	٣٣,٧	٣٩,١	٤٢,٦	٤٥,٧	٤٩,٦
٣٠	١٥,٠	١٦,٨	٢٠,٦	٢٤,٥	٢٩,٣	٣٤,٨	٤٠,٣	٤٣,٨	٤٧,٠	٥٠,٩
٤٠	٢٢,٢	٢٤,٤	٢٩,١	٣٣,٧	٣٩,٣	٤٥,٦	٥١,٨	٥٥,٨	٥٩,٣	٦٣,٧
٥٠	٢٩,٧	٣٢,٤	٣٧,٧	٤٢,٩	٤٩,٣	٥٦,٣	٦٣,٢	٦٧,٥	٧١,٤	٧٦,٢
٦٠	٣٧,٥	٤٠,٥	٤٦,٥	٥٢,٣	٥٩,٣	٦٧,٠	٧٤,٤	٧٩,١	٨٣,٣	٨٨,٤

جدول رقم (٢ - ١١)

القيم الحرجة لاختبار كولموجروف - سميرنوف " د "

(اختبار "د")

درجات الحرية	مستوى الدلالة				
	٠٫٢٠	٠٫١٥	٠٫١٠	٠٫٠٥	٠٫٠١
١	٠٫٩٠٠	٠٫٨٢٥	٠٫٨٥٠	٠٫٩٧٥	٠٫٩٩٥
٢	٠٫٦٨٤	٠٫٧٢٦	٠٫٧٧٦	٠٫٨٤٢	٠٫٩٢٩
٣	٠٫٥٦٥	٠٫٥٩٧	٠٫٦٤٢	٠٫٧٠٣	٠٫٨٢٩
٤	٠٫٤٩٤	٠٫٥٢٦	٠٫٥٦٤	٠٫٦٢٤	٠٫٧٣٤
٥	٠٫٤٤٦	٠٫٤٧٤	٠٫٥١٠	٠٫٥٦٤	٠٫٦٦٩
٦	٠٫٤١٠	٠٫٤٣٦	٠٫٤٧٠	٠٫٥٢١	٠٫٦١٨
٧	٠٫٣٨١	٠٫٤٠٥	٠٫٤٣٨	٠٫٤٨٦	٠٫٥٧٧
٨	٠٫٣٥٨	٠٫٣٨١	٠٫٤١١	٠٫٤٥٤	٠٫٥٤٣
٩	٠٫٣٣٩	٠٫٣٦٠	٠٫٣٨٥	٠٫٤٣٢	٠٫٥١٤
١٠	٠٫٣٢٢	٠٫٣٤٢	٠٫٣٦٨	٠٫٤٠٩	٠٫٤٨٦
١١	٠٫٣٠٧	٠٫٣٢٦	٠٫٣٥٢	٠٫٣٩١	٠٫٤٦٨
١٢	٠٫٢٩٥	٠٫٣١٣	٠٫٣٣٨	٠٫٣٧٥	٠٫٤٥٠
١٣	٠٫٢٨٤	٠٫٣٠٢	٠٫٣٢٥	٠٫٣٦١	٠٫٤٣٣
١٤	٠٫٢٧٤	٠٫٢٩٢	٠٫٣١٤	٠٫٣٤٩	٠٫٤١٨
١٥	٠٫٢٦٦	٠٫٢٨٣	٠٫٣٠٤	٠٫٣٣٨	٠٫٤٠٤
١٦	٠٫٢٥٨	٠٫٢٧٤	٠٫٢٩٥	٠٫٣٢٨	٠٫٣٩١
١٧	٠٫٢٥٠	٠٫٢٦٦	٠٫٢٨٦	٠٫٣١٥	٠٫٣٨٠
١٨	٠٫٢٤٤	٠٫٢٥٩	٠٫٢٧٨	٠٫٣٠٩	٠٫٣٧٠
١٩	٠٫٢٣٧	٠٫٢٥٢	٠٫٢٧٢	٠٫٣٠١	٠٫٣٦١
٢٠	٠٫٢٣١	٠٫٢٤٦	٠٫٢٦٥	٠٫٢٩١	٠٫٣٥٢
٢٥	٠٫٢١	٠٫٢٢	٠٫٢٤	٠٫٢٦٤	٠٫٣٢
٣٠	٠٫١٩	٠٫٢٠	٠٫٢٢	٠٫٢٤٢	٠٫٢٩
٣٥	٠٫١٨	٠٫١٩	٠٫٢١	٠٫٢٣	٠٫٢٧
٤٠	٠٫١٧	٠٫١٨	٠٫١٩	٠٫٢١	٠٫٢٥
٥٠	٠٫١٥	٠٫١٦	٠٫١٧	٠٫١٩	٠٫٢٣
٦٠	٠٫١٤	٠٫١٥	٠٫١٦	٠٫١٧	٠٫٢١
٧٠	٠٫١٣	٠٫١٤	٠٫١٥	٠٫١٦	٠٫١٩
٨٠	٠٫١٢	٠٫١٣	٠٫١٤	٠٫١٥	٠٫١٨
٩٠	٠٫١١	٠٫١٢	٠٫١٣	٠٫١٤	٠٫١٧
١٠٠	٠٫١١	٠٫١١	٠٫١٢	٠٫١٤	٠٫١٦
أكثر من ١٠٠	$\frac{١٫٠٧}{\sqrt{ن}}$	$\frac{١٫١٤}{\sqrt{ن}}$	$\frac{١٫٢٢}{\sqrt{ن}}$	$\frac{١٫٣٦}{\sqrt{ن}}$	$\frac{١٫٦٣}{\sqrt{ن}}$

(يرفض فرض العدم عندما تكون قيمة "د" أكبر من القيمة الحرجة عند مستوى الدلالة المطلوب).

جدول رقم (٣-١٢)

القيم الحرجة لاختبار مان - هويتني (اختبار ي) عند مستوى دلالة ٠,٠١ [اختبار ثنائي الطرف]

ن٢ \ ن١	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	١١	١٢	١٣	١٤	١٥	١٦	١٧	١٨	١٩	٢٠
١																			٠	٠
٢					٠	٠	٠	١	١	١	١	٢	٢	٢	٣	٣	٣	٤	٤	٤
٣				٠	١	٢	٢	٣	٣	٤	٥	٥	٦	٧	٧	٨	٩	٩	١٠	١١
٤				١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	١١	١٢	١٤	١٥	١٦	١٧	١٨
٥		٠	١	٢	٤	٥	٦	٦	٩	١١	١٢	١٣	١٥	١٦	١٨	١٩	٢٠	٢٢	٢٣	٢٥
٦		٠	٢	٣	٥	٧	٨	١٠	١٢	١٤	١٦	١٧	١٩	٢١	٢٣	٢٥	٢٦	٢٨	٣٠	٣٢
٧		٠	٢	٤	٦	٨	١١	١٣	١٦	١٧	١٩	٢١	٢٤	٢٦	٢٨	٣٠	٣٣	٣٥	٣٧	٣٩
٨		١	٣	٥	٨	١٠	١٣	١٥	١٨	٢٠	٢٣	٢٦	٢٨	٣١	٣٣	٣٦	٣٩	٤١	٤٤	٤٣
٩		١	٣	٦	٩	١٢	١٥	١٨	٢١	٢٤	٢٧	٣٠	٣٣	٣٦	٣٩	٤٢	٤٥	٤٨	٥١	٥٤
١٠		١	٤	٧	١١	١٤	١٧	٢٠	٢٤	٢٧	٣١	٣٤	٣٧	٤١	٤٤	٤٨	٥١	٥٥	٥٨	٦٢
١١		١	٥	٨	١٢	١٦	١٩	٢٣	٢٧	٣١	٣٤	٣٨	٤٢	٤٦	٥٠	٥٤	٥٧	٦١	٦٥	٦٩
١٢		٢	٥	٩	١٣	١٧	٢١	٢٦	٣٠	٣٤	٣٨	٤٢	٤٧	٥١	٥٥	٦٠	٦١	٦٨	٧٢	٧٧
١٣		٢	٦	١٠	١٥	١٩	٢٤	٢٨	٣٣	٣٧	٤٢	٤٧	٥١	٥٦	٦١	٦٥	٧٠	٧٥	٨٠	٨٤
١٤		٢	٧	١١	١٦	٢١	٢٦	٣١	٣٦	٤١	٤٦	٥١	٥٦	٦١	٦٦	٧١	٧٧	٨٢	٨٧	٩٢
١٥		٣	٧	١٢	١٨	٢٣	٢٨	٣٣	٣٩	٤٤	٥٠	٥٥	٦١	٦٦	٧٢	٧٧	٨٢	٨٨	٩١	١٠٠
١٦		٣	٨	١٤	١٩	٢٥	٣٠	٣٦	٤٢	٤٨	٥٤	٦٠	٦٥	٧١	٧٧	٨٣	٨٩	٩٥	١٠١	١٠٧
١٧		٣	٩	١٥	٢٠	٢٦	٣٣	٣٩	٤٥	٥١	٥٧	٦٤	٧٠	٧٧	٨٣	٨٩	٩٦	١٠٢	١٠٩	١١٥
١٨		٤	٩	١٦	٢٢	٢٨	٣٥	٤١	٤٨	٥٥	٦١	٦٨	٧٥	٨٢	٨٨	٩٥	١٠٢	١٠٩	١١٦	١٢٣
١٩	٠	٤	١٠	١٧	٢٣	٣٠	٣٧	٤٤	٥١	٥٨	٦٥	٧٢	٨٠	٨٧	٩٢	١٠١	١٠٩	١١٦	١٢٣	١٣٠
٢٠	٠	٤	١١	١٨	٢٥	٣٢	٣٩	٤٧	٥٤	٦٢	٦٩	٧٧	٨٤	٩٢	١٠٠	١٠٧	١١٥	١٢٣	١٣٠	١٣٢

(يرفض فرض العدم عندما تكون قيمة "ي" المحسوبة أقل من أو تساوي القيمة الحرجة عند مستوى الدلالة المطلوب)

(تابع) جدول رقم (٣-١٠)

القيم الحرجة لاختبار مان - هويتني (اختبار - ى) عند مستوى دلالة ٠.٠٢ [اختبار ثنائي الطرف]

ن٢ / ن١	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	١١	١٢	١٣	١٤	١٥	١٦	١٧	١٨	١٩	٢٠
١																				
٢													٠	٠	٠	٠	٠	٠	١	١
٣							٠	١	١	١	١	٢	٢	٢	٣	٣	٤	٤	٤	٥
٤					٠	١	١	٢	٣	٢	٤	٥	٥	٦	٧	٧	٨	٩	٩	١٠
٥				٠	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	١١	١٢	١٣	١٤	١٥	١٦
٦				١	٢	٣	٤	٦	٧	٨	٩	١١	١٢	١٣	١٥	١٦	١٨	١٩	٢٠	٢٢
٧			٠	١	٣	٤	٦	٨	٩	١١	١٢	١٤	١٦	١٧	١٩	٢١	٢٣	٢٤	٢٦	٢٨
٨			١	٢	٤	٦	٨	١٠	١١	١٣	١٥	١٧	٢٠	٢٢	٢٤	٢٦	٢٨	٣٠	٣٢	٣٤
٩			١	٣	٥	٧	٩	١١	١٤	١٦	١٨	٢١	٢٣	٢٦	٢٨	٣١	٣٣	٣٦	٣٨	٤٠
١٠			١	٣	٦	٨	١١	١٣	١٦	١٩	٢٢	٢٤	٢٧	٣٠	٣٣	٣٦	٣٨	٤١	٤٤	٤٧
١١			١	٤	٧	٩	١٢	١٥	١٨	٢٢	٢٥	٢٨	٣١	٣٤	٣٧	٤١	٤٤	٤٧	٥٠	٥٣
١٢			٢	٥	٨	١١	١٤	١٧	٢١	٢٤	٢٨	٣١	٣٥	٣٨	٤٢	٤٦	٤٩	٥٣	٥٦	٦٠
١٣		٠	٢	٥	٩	١٢	١٦	٢٠	٢٣	٢٧	٣١	٣٥	٣٩	٤٣	٤٧	٥١	٥٥	٥٩	٦٣	٦٧
١٤		٠	٢	٦	١٠	١٣	١٧	٢٢	٢٦	٣٠	٣٤	٣٨	٤٣	٤٧	٥١	٥٦	٦٠	٦٥	٦٩	٧٣
١٥		٠	٣	٧	١١	١٥	١٩	٢٤	٢٨	٣٣	٣٧	٤٢	٤٧	٥١	٥٦	٦١	٦٦	٧٠	٧٥	٨٠
١٦		٠	٣	٧	١٢	١٦	٢١	٢٦	٣١	٣٦	٤١	٤٦	٥١	٥٦	٦١	٦٦	٧١	٧٦	٨٢	٨٧
١٧		٠	٤	٨	١٣	١٨	٢٣	٢٨	٣٣	٣٨	٤٤	٤٩	٥٥	٦٠	٦٦	٧١	٧٧	٨٢	٨٨	٩٣
١٨		٠	٤	٩	١٤	١٩	٢٤	٣٠	٣٦	٤١	٤٧	٥٣	٥٩	٦٥	٧٠	٧٦	٨٢	٨٨	٩٤	١٠٠
١٩		١	٤	٩	١٥	٢٠	٢٦	٣٢	٣٨	٤٤	٥٠	٥٦	٦٣	٦٩	٧٥	٨٢	٨٨	٩٤	١٠١	١٠٧
٢٠		١	٥	١٠	١٦	٢٢	٢٨	٣٤	٤٠	٤٧	٥٣	٦٠	٦٧	٧٣	٨٠	٨٧	٩٣	١٠٠	١٠٧	١١٤

(تابع) جدول رقم (٣ - ١٢)

القيم الحرجة لاختبار مان - هويتنى (اختبار ى) عند مستوى دلالة ٠٫٠٥ [اختبار ثنائى الطرف]

ن١ / ن٢	١	١	٢	٣	٤	٥	٧	٨	٩	١٠	١١	١٢	١٣	١٤	١٥	١٦	١٧	١٨	١٩	٢٠
١	٠																			
٢	٠							٠	٠	٠	٠	١	١	١	١	١	٢	٢	٢	٢
٣	٠				٠	١	١	٢	٢	٣	٣	٤	٤	٥	٥	٦	٦	٧	٧	٨
٤	٠			٠	١	٢	٣	٤	٤	٥	٦	٧	٨	٩	١٠	١١	١١	١٢	١٣	١٣
٥	٠		٠	١	٢	٣	٥	٦	٧	٨	٩	١١	١٢	١٣	١٤	١٥	١٧	١٨	١٩	٢٠
٦	٠		١	٢	٣	٥	٦	٨	١٠	١١	١٣	١٤	١٦	١٧	١٩	٢١	٢٢	٢٤	٢٥	٢٧
٧	٠		١	٣	٥	٦	٨	١٠	١٢	١٤	١٦	١٨	٢٠	٢٢	٢٤	٢٦	٢٨	٣٠	٣٢	٣٤
٨	٠	٠	٢	٤	٦	٨	١٠	١٣	١٥	١٧	١٩	٢٢	٢٤	٢٦	٢٩	٣١	٣٤	٣٦	٣٨	٤١
٩		٠	٢	٤	٧	١٠	١٢	١٥	١٧	٢٠	٢٣	٢٦	٢٨	٣١	٣٤	٣٧	٣٩	٤٢	٤٥	٤٨
١٠	٠	٠	٣	٥	٨	١١	١٤	١٧	٢٠	٢٣	٢٦	٢٩	٣٣	٣٦	٣٩	٤٢	٤٥	٤٨	٥٢	٥٥
١١	٠	٠	٣	٦	٩	١٣	١٦	١٩	٢٣	٢٦	٣٠	٣٣	٣٧	٤٠	٤٤	٤٧	٥١	٥٥	٥٨	٦٢
١٢	٠	١	٤	٧	١١	١٤	١٨	٢٢	٢٦	٢٩	٣٣	٣٧	٤١	٤٥	٤٩	٥٣	٥٧	٦١	٦٥	٦٩
١٣	٠	١	٤	٨	١٢	١٦	٢٠	٢٤	٢٨	٣٣	٣٧	٤١	٤٥	٥٠	٥٤	٥٩	٦٣	٦٧	٧٢	٧٦
١٤	٠	١	٥	٩	١٣	١٧	٢٢	٢٦	٣١	٣٦	٤٠	٤٥	٥٠	٥٥	٥٩	٦٤	٧٦	٧٤	٧٨	٨٣
١٥	٠	١	٥	١٠	١٤	١٩	٢٤	٢٩	٣٤	٣٩	٤٤	٤٦	٥٤	٥٩	٦٤	٧٠	٧٥	٨٠	٨٥	٩٠
١٦	٠	١	٦	١١	١٥	٢١	٢٦	٣١	٣٧	٤٢	٤٧	٥٣	٥٩	٦٤	٧٠	٧٥	٨١	٨٦	٩٢	٩٨
١٧	٠	٢	٦	١١	١٧	٢٢	٢٨	٣٤	٣٩	٤٥	٥١	٥٧	٦٣	٦٧	٧٥	٨١	٨٧	٩٣	٩٩	١٠٥
١٨	٠	٢	٧	١٢	١٨	٢٤	٣٠	٣٦	٤٢	٤٨	٥٥	٦١	٦٧	٧٤	٨٠	٨٦	٩٣	٩٩	١٠٦	١١٢
١٩	٠	٢	٧	١٣	١٩	٢٥	٣٢	٣٨	٤٥	٥٢	٥٨	٦٥	٧٢	٧٨	٨٥	٩٢	٩٩	١٠٦	١١٣	١١٩
٢٠	٠	٢	٨	١٣	٢٠	٢٧	٣٤	٤١	٤٨	٥٥	٦٢	٦٩	٧٦	٨٣	٩٠	٩٨	١٠٥	١١٢	١١٩	١٢٧

جدول رقم (م - ١٢)

القيم الحرجة لاختبار ويلكوكسون (اختبار - ت)

عدد أزواج القيم = ن	أ ٠٥ر / ب ٠٢٥ر (٢ - ٠٥ر)	٠٢ر / ٠١ر	٠١ر / ٠٠٥ر
٦	٠	-	-
٧	٢	٠	-
٨	٤	٢	٠
٩	٦	٣	٢
١٠	٨	٥	٣
١١	١١	٧	٥
١٢	١٤	١٠	٧
١٣	١٧	١٣	١٠
١٤	٢١	١٦	١٣
١٥	٢٥	٢٠	١٦
١٦	٣٠	٢٤	٢٠
١٧	٣٥	٢٨	٢٣
١٨	٤٠	٣٣	٢٨
١٩	٤٦	٣٨	٣٢
٢٠	٥٢	٤٣	٣٨
٢١	٥٩	٤٩	٤٣
٢٢	٦٦	٥٦	٤٩
٢٣	٧٣	٦٢	٥٥
٢٤	٨١	٦٩	٦١
٢٥	٨٩	٧٧	٦٨

أ = اختبار ثنائي الطرف

ب = اختبار أحادي الطرف

(قيمة ت المحسوبة تكون ذات دلالة إحصائية إذا كانت أقل من القيمة النظرية لها في الجدول عند مستوى الدلالة المطلوب).

جدول رقم (٣ - ١٤)

القيم الحرجة لاختبار كروسكال - واليس

اختبار H .

ن١	ن٢	ن٣	مستوى الدلالة ١٠ر	٠٥ر	٠١ر	٠٠٥ر
٢	١	١				
٢	٢	١				
٢	٢	٢	٤,٥٧١			
٣	١	١				
٣	٢	١	٤,٢٨٦			
٣	٢	٢	٤,٥٠٠	٤,٧١٤	٥,٣٥٧	
٣	٣	١	٤,٥٧١	٥,١٤٣		
٣	٣	٢	٤,٥٥٦	٥,٣٦١		
٣	٣	٣	٤,٦٢٢	٥,٦٠٠	٧,٢٠٠	٧,٢٠٠
٤	١	١				
٤	٢	١	٤,٥٠٠			
٤	٢	٢	٤,٠٥٦	٥,٢٠٨		
٤	٣	٢	٤,٥١١	٥,٤٤٤	٦,٤٤٤	
٤	٣	٣	٤,٧٠٩	٥,٧٢٧	٦,٧٤٦	
٤	٤	١	٤,١٦٧	٤,٩٦٧	٦,٦٦٧	
٤	٤	٢	٤,٥٥٥	٥,٤٥٥	٧,٠٣٦	
٤	٤	٣	٤,٥٤٦	٥,٥٩٩	٧,١٤٤	
٤	٤	٤	٤,٦٥٤	٥,٦٩٢	٧,٦٥٤	
٥	١	١				
٥	٢	١	٤,٢٠٠	٥,٠٠٠		
٥	٢	٢	٤,٣٧٣	٥,١٦٠	٦,٥٣٣	
٥	٣	١	٤,٠١٨	٤,٩٦٠		
٥	٣	٢	٤,٦٥١	٥,٢٥١	٦,٨٢٢	
٥	٣	٣	٤,٥٣٣	٤,٦٤٩	٧,٠٧٩	
٥	٤	١	٣,٩٨٧	٤,٩٨٦	٦,٩٥٥	
٥	٤	٢	٤,٥٤١	٥,٢٦٨	٧,١١٨	
٥	٤	٣	٤,٥٤٩	٥,٦٣١	٧,٤٤٥	
٥	٤	٤	٤,٦١٩	٥,٦١٨	٧,٧٦٠	
٥	٥	١	٤,١٠٩	٥,١٢٧	٧,٣٠٩	
٥	٥	٢	٤,٥٠٨	٥,٣٣٩	٧,٢٦٩	
٥	٥	٣	٤,٥٤٥	٥,٧٠٦	٧,٥٤٣	
٥	٥	٤	٤,٥٢٣	٥,٦٤٣	٧,٧٩١	
٥	٥	٥	٤,٥٦٠	٥,٧٨٠	٧,٩٨٠	

(يرفض فرض العدم عندما تكون قيمة H أكبر من أو تساوي القيمة الحرجة عند مستوى الدلالة المطلوب) .

جدول رقم (م-١٥): قيمة أقل معامل ارتباط معنوي عند مستوى دلالة ٠٫٠٥ ، ٠٫٠١

درجات الحرية	٥	١	درجات الحرية	٥	١
١	٠٫٩٩٧	١٫٠٠٠	٢٤	٠٫٣٨٨	٠٫٤٩٦
٢	٠٫٩٥٠	٠٫٩٩٠	٢٥	٠٫٣٨١	٠٫٤٨٧
٣	٠٫٨٧٨	٠٫٩٥٩	٢٦	٠٫٣٧٤	٠٫٤٧٨
٤	٠٫٨١١	٠٫٩١٧	٢٧	٠٫٣٦٧	٠٫٤٧٠
٥	٠٫٧٥٤	٠٫٨٧٤	٢٨	٠٫٣٦١	٠٫٤٦٣
٦	٠٫٧٠٧	٠٫٨٣٤	٢٩	٠٫٣٥٥	٠٫٤٥٦
٧	٠٫٦٦٦	٠٫٧٩٨	٣٠	٠٫٣٤٩	٠٫٤٤٩
٨	٠٫٦٣٢	٠٫٧٦٥	٣٥	٠٫٣٢٥	٠٫٤١٨
٩	٠٫٦٠٢	٠٫٧٣٥	٤٠	٠٫٣٠٤	٠٫٣٩٣
١٠	٠٫٥٧٦	٠٫٧٠٨	٤٥	٠٫٢٨٨	٠٫٣٧٢
١١	٠٫٥٥٣	٠٫٦٨٤	٥٠	٠٫٢٧٣	٠٫٣٥٤
١٢	٠٫٥٣٢	٠٫٦٦١	٦٠	٠٫٢٥٠	٠٫٣٢٥
١٣	٠٫٥١٤	٠٫٦٤١	٧٠	٠٫٢٣٢	٠٫٣٠٢
١٤	٠٫٤٩٧	٠٫٦٢٣	٨٠	٠٫٢١٧	٠٫٢٨٣
١٥	٠٫٤٨٢	٠٫٦٠٦	٩٠	٠٫٢٠٥	٠٫٢٦٧
١٦	٠٫٤٦٨	٠٫٥٩٠	١٠٠	٠٫١٩٥	٠٫٢٥٤
١٧	٠٫٤٥٦	٠٫٥٧٥	١٢٥	٠٫١٧٤	٠٫٢٢٨
١٨	٠٫٤٤٤	٠٫٥٦١	١٥٠	٠٫١٥٩	٠٫٢٠٨
١٩	٠٫٤٣٣	٠٫٥٤٩	٢٠٠	٠٫١٣٨	٠٫١٨١
٢٠	٠٫٤٢٣	٠٫٥٣٧	٣٠٠	٠٫١١٣	٠٫١٤٨
٢١	٠٫٤١٣	٠٫٥٢٦	٤٠٠	٠٫٠٩٨	٠٫١٢٨
٢٢	٠٫٤٠٤	٠٫٥١٥	٥٠٠	٠٫٠٨٨	٠٫١١٥
٢٣	٠٫٣٩٦	٠٫٥٠٥	١٠٠٠	٠٫٠٦٢	٠٫٠٨١

جدول (م-١٦)

قيمة أقل معامل ارتباط معنوي لمعامل ارتباط الرتب (سبيرمان)

مستويات الدلالة / درجات الحرية	أ- ٠٫١٠ / ب- ٠٫٠٥	٠٫٠٥ / ٠٫٠٢٥	٠٫٠٢ / ٠٫٠١	٠٫٠١ / ٠٫٠٠٥
٤	١٫٠٠٠			
٥	٠٫٩٠٠	١٫٠٠٠	١٫٠٠٠	
٦	٠٫٨٢٩	٠٫٨٨٦	٠٫٩٤٣	١٫٠٠٠
٧	٠٫٧١٤	٠٫٧٨٦	٠٫٨٩٣	٠٫٩٢٩
٨	٠٫٦٤٣	٠٫٧٣٨	٠٫٨٣٣	٠٫٨٨١
٩	٠٫٦٠٠	٠٫٦٨٣	٠٫٧٨٣	٠٫٨٣٣
١٠	٠٫٥٦٤	٠٫٦٤٨	٠٫٧٤٥	٠٫٧٩٤
١١	٠٫٥٢٣	٠٫٦٢٣	٠٫٧٣٦	٠٫٨١٨
١٢	٠٫٤٩٧	٠٫٥٩١	٠٫٧٠٣	٠٫٧٨٠
١٣	٠٫٤٧٥	٠٫٥٦٦	٠٫٦٧٣	٠٫٧٤٥
١٤	٠٫٤٥٧	٠٫٥٤٥	٠٫٦٤٦	٠٫٧١٦
١٥	٠٫٤٤١	٠٫٥٢٥	٠٫٦٢٣	٠٫٦٨٩
١٦	٠٫٤٢٥	٠٫٥٠٧	٠٫٦٠١	٠٫٦٦٦
١٧	٠٫٤١٢	٠٫٤٩٠	٠٫٥٨٢	٠٫٦٤٥
١٨	٠٫٣٩٩	٠٫٤٧٦	٠٫٥٦٤	٠٫٦٢٥
١٩	٠٫٣٨٨	٠٫٤٦٢	٠٫٥٤٩	٠٫٦٠٨
٢٠	٠٫٣٧٧	٠٫٤٥٠	٠٫٥٣٤	٠٫٥٩١
٢١	٠٫٣٦٨	٠٫٤٣٨	٠٫٥٢١	٠٫٥٧٦
٢٢	٠٫٣٥٩	٠٫٤٢٨	٠٫٥٠٨	٠٫٥٦٢
٢٣	٠٫٣٥١	٠٫٤١٨	٠٫٤٩٦	٠٫٥٤٩
٢٤	٠٫٣٤٣	٠٫٤٠٩	٠٫٤٨٥	٠٫٥٣٧
٢٥	٠٫٣٣٦	٠٫٤٠٠	٠٫٤٧٥	٠٫٥٢٦
٢٦	٠٫٣٢٩	٠٫٣٩٢	٠٫٤٦٥	٠٫٥١٥
٢٧	٠٫٣٢٣	٠٫٣٨٥	٠٫٤٥٦	٠٫٥٠٥
٢٨	٠٫٣١٧	٠٫٣٧٧	٠٫٤٤٨	٠٫٤٩٦
٢٩	٠٫٣١١	٠٫٣٧٠	٠٫٤٤٠	٠٫٤٨٧
٣٠	٠٫٣٠٥	٠٫٣٦٤	٠٫٤٣٢	٠٫٤٧٨
٣٥	٠٫٢٨٢	٠٫٣٣٦	٠٫٣٩٩	٠٫٤٤٢
٤٠	٠٫٢٦٣	٠٫٣١٤	٠٫٣٧٣	٠٫٤١٣
٤٥	٠٫٢٤٨	٠٫٢٩٦	٠٫٣٥١	٠٫٣٨٨
٥٠	٠٫٢٣٥	٠٫٢٨٠	٠٫٣٣٢	٠٫٣٦٨
٥٥	٠٫٢٢٤	٠٫٢٦٧	٠٫٣١٧	٠٫٣٥١
٦٠	٠٫٢١٤	٠٫٢٥٥	٠٫٣٠٣	٠٫٣٣٥
٦٥	٠٫٢٠٦	٠٫٢٤٥	٠٫٢٩١	٠٫٣٢٢
٧٠	٠٫١٩٨	٠٫٢٣٦	٠٫٢٨٠	٠٫٣١٠
٧٥	٠٫١٩١	٠٫٢٢٨	٠٫٢٧١	٠٫٣٠٠
٨٠	٠٫١٨٥	٠٫٢٢١	٠٫٢٦٢	٠٫٢٩٠
٨٥	٠٫١٨٠	٠٫٢١٤	٠٫٢٥٤	٠٫٢٨١
٩٠	٠٫١٧٤	٠٫٢٠٨	٠٫٢٤٧	٠٫٢٧٣
٩٥	٠٫١٧٠	٠٫٢٠٢	٠٫٢٤٠	٠٫٢٦٦
١٠٠	٠٫١٦٥	٠٫١٩٧	٠٫٢٣٤	٠٫٢٥٩

أ- اختبار ثنائي الطرف   ب- اختبار أحادي الطرف

جدول رقم (٣ - ١٧)

القيم الاحتمالية لحدوث معامل ارتباط كندال للرتب بفعل الصدفة أو المتغيرات العشوائية

س	عدد ازواج القيم = ن			
	٤	٥	٨	٩
٠	٠٫٦٢٥	٠٫٥٩٢	٠٫٥٤٨	٠٫٥٤٠
٢	٠٫٣٧٥	٠٫٤٠٨	٠٫٤٥٢	٠٫٤٦٠
٤	٠٫١٦٧	٠٫٢٤٢	٠٫٣٦٠	٠٫٣٨١
٦	٠٫٠٤٢	٠٫١١٧	٠٫٢٧٤	٠٫٣٠٦
٨		٠٫٠٤٢	٠٫١٩٩	٠٫٢٣٨
١٠		٠٫٠٠٨٣	٠٫١٣٨	٠٫١٧٩
١٢			٠٫٠٨٩	٠٫١٣٠
١٤			٠٫٠٥٤	٠٫٠٩٠
١٦			٠٫٠٣١	٠٫٠٦٠
١٨			٠٫٠١٦	٠٫٠٣٨
٢٠			٠٫٠٠٧١	٠٫٠٢٢
٢٢			٠٫٠٠٢٨	٠٫٠١٢
٢٤			٠٫٠٠٠٨٣	٠٫٠٠٦٣
٢٦			٠٫٠٠٠١٩	٠٫٠٠٢٩
٢٨			٠٫٠٠٠٠٢٥	٠٫٠٠١٢

س	عدد ازواج القيم = ن		
	٦	٧	١٠
١	٠٫٥٠٠	٠٫٥٠٠	٠٫٥٠٠
٣	٠٫٣٦٠	٠٫٣٨٦	٠٫٤٣١
٥	٠٫٢٣٥	٠٫٢٨١	٠٫٣٦٤
٧	٠٫١٣٦	٠٫١٩١	٠٫٣٠٠
٩	٠٫٠٦٨	٠٫١١٩	٠٫٢٤٢
١١	٠٫٠٢٨	٠٫٠٦٨	٠٫١٩٠
١٣	٠٫٠٠٨٣	٠٫٠٣٥	٠٫١٤٦
١٥	٠٫٠٠١٤	٠٫٠١٥	٠٫١٠٨
١٧		٠٫٠٠٥٤	٠٫٠٧٨
١٩		٠٫٠٠١٤	٠٫٠٥٤
٢١		٠٫٠٠٠٢٠	٠٫٠٣٦
٢٣			٠٫٠٢٨
٢٥			٠٫٠١٤
٢٧			٠٫٠٠٨٣
٢٩			٠٫٠٠٤٦
٣١			٠٫٠٠٢٣
٣٣			٠٫٠٠١١



